Environmental shaping of bacterial communities and functional gene diversity in the extreme Transantarctic Mountains

Joana Séneca Cardoso da Silva

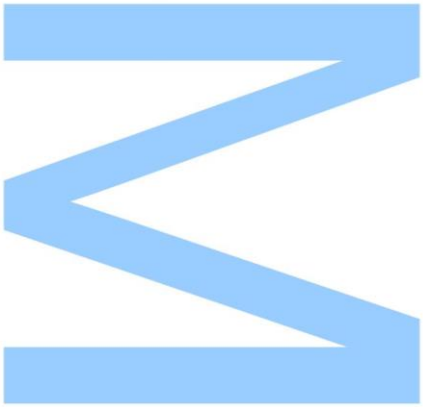
Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto em
Biodiversidade, Genética e Evolução
2013/2014

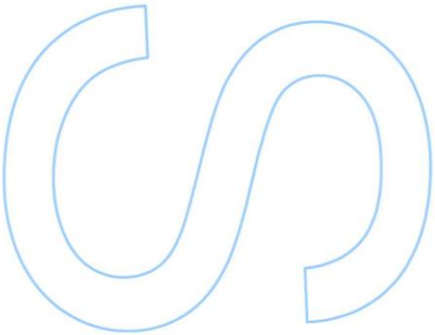
Orientadora
Catarina Maria Pinto Mora Pinto de Magalhães
PhD, Investigadora no Centro Interdisciplinar de Investigação Marinha e Ambiental (CIIMAR)

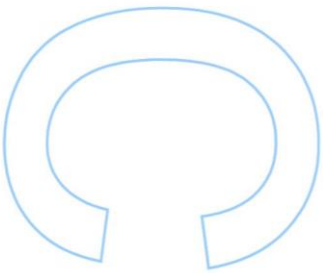
Co-orientador
Stephen Craig Cary
PhD, Professor na Universidade de Waikato, Nova Zelândia

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____







Dissertação de candidatura ao grau de Mestre em Biodiversidade, Genética e Evolução submetida à Faculdade de Ciências da Universidade do Porto.

O presente trabalho foi desenvolvido no Centro Interdisciplinar de Investigação Marinha e Ambiental (CIIMAR), sob a orientação científica da Doutora Catarina Maria Pinto Mora Pinto de Magalhães e do Doutor Stephen C. Cary

Dissertation for applying to a Master's degree in Biodiversity, Genetics and Evolution, submitted to the Faculty of Sciences of the University of Porto.

The present work was developed at the Interdisciplinary Center for Marine and Environmental Research (CIIMAR)
under the scientific supervision of Doutora Catarina Maria Pinto Mora Pinto de Magalhães and Doctor Stephen C. Cary.

“Somewhere, something incredible is waiting to be known”

Carl Sagan

Agradecimentos

Antes de mais,

Um grande obrigada à minha orientadora, Doutora Catarina Magalhães, pelo constante apoio, disponibilidade e entusiasmo com que me recebeu e orientou ao longo deste ano. Um agradecimento também ao Prof. Dr. Anthony Danko por ter tão amavelmente disponibilizado os equipamentos necessários, na FEUP.

Aos meus colegas EcoBioTec, em particular ao Hugo Ribeiro pelos ensinamentos iniciais e momentos de companheirismo (e por me tirar as amostras dos mais variados equipamentos, durante o seu turno da noite). À imprescindível Maria Monteiro que aturou e partilhou as minhas frustrações, sem nunca abrir mão. À Paula Salgado, pela boa disposição constante e apoio, profissional e pessoalmente. À Ana Paula Mucha e à Sandra Ramos pela ajuda e conhecimento nas tarefas laboratoriais do dia-a-dia. Aos meus colegas de mestrado neste laboratório Joana, Marisa e Nuno, pelo apoio e momentos de convívo.

À turma MBGE 2012/2013 por todos os momentos, trocas de opinião, esclarecimento de pequenas dúvidas aparentemente óbvias e ridículas, e apoio mútuo, que foram absolutamente imprescindíveis para a conclusão deste trabalho. Um particular obrigado à Marina por ter sido a minha dupla na maioria dos relatórios, por compreender o dilema do computador lento, e por ter estado sempre disponível para debater dúvidas existenciais. Yes, we made it!

A todos os meus amigos e colegas que directa ou indirectamente, estando longe ou perto, me ajudaram ao longo deste ano.

Ao César por estar sempre, sempre disposto a ouvir-me.

Um grande obrigada também dedicado à minha família pelo apoio incondicional. Em particular à minha mãe, por ser (e ter sido sempre) o meu principal ponto de apoio.

Agradeço também à Fundação Portuguesa de Tecnologia (FCT) pelo financiamento do projecto NITROEXTREM (PTDC/MAR/112723/2009– FCOMP-01-0124-FEDER-015422), à equipa NZTABS/ICTAR, e ao Programa Polar Português.

Resumo

As regiões livres de gelo na Antártida representam apenas 0.3 % de toda a área do continente, e estão distribuídas de forma heterogénea. As extremas condições ambientais existentes nestas localizações moldaram ecossistemas caracterizados por uma estrutura trófica simples, nos quais os microrganismos são dominantes e enfrentam severas condições, como uma baixíssima disponibilidade de água e nutrientes, temperaturas baixas, ciclos de congelamento/descongelamento, longos períodos de escuridão no Inverno, e exposição a altos níveis de radiação ultravioleta no Verão. Nestes ecossistemas extremos, a diversidade microbiana geral e funcional permanece largamente desconhecida.

Este estudo faz parte de um programa multidisciplinar liderado pela Nova Zelândia (NZTABS), cujo foco primário é a caracterização da diversidade microbiana nos vales secos da Antártica. Na expedição K020 ao Vale Vitória em Janeiro de 2013 amostrou-se uma área correspondente a 300 km^2. Foram recolhidas amostras de solo, e as condições ambientais de cada ponto de amostragem (total 86) foram registadas *in situ* (metadata). As comunidades microbianas destes locais foram também caracterizadas ao nível da actividade, abundância e diversidade, através da medição dos níveis de ATP, através da coloração por DAPI, e também pelo método de análise da região intergénica bacteriana 16S-23S DNA (ARISA). Relações entre os parâmetros ambientais e biológicos recolhidos nas 86 estações de amostragem revelaram que a diversidade, abundância e actividade das comunidades microbianas do Vale Vitória estão sujeitas a uma forte estruturação espacial, resultante da heterogeneidade dos parâmetros físico-químicos que caracterizam estes ambientes. De entre as variáveis ambientais analisadas, e de acordo com estudos prévios nestas áreas, a disponibilidade de água foi identificada como um dos principais factores limitantes à distribuição das comunidades microbianas.

De modo a compreender com maior detalhe de que forma a disponibilidade de água afecta as comunidades microbianas nestes ambientes extremos, foram realizadas amostragens ao longo de um transecto com um gradiente de disponibilidade de água no Vale Vitoria. Dos seis pontos do transecto foi extraído DNA e o gene 16S rRNA foi pirosequenciado usando a tecnologia da Roche (454). Métricas de diversidade (alfa e beta) foram produzidas para cada ponto de amostragem e contrastadas com os parâmetros ambientais recolhidos nos locais. Os resultados revelaram uma clara mudança nas comunidades microbianas existentes, nomeadamente uma substituição notória entre os filos Proteobacteria e Actinobacteria, à medida que os solos iam ficando mais secos. Adicionalmente, o local com maior número de OTUs correspondeu à frequência máxima de membros do filo Bacteroidetes

e também ao local com maiores percentagens de carbono orgânico e conductividade. Membros do filo Cianobacteria diminuíram a sua frequência à medida que os solos se tornaram mais secos, o que possivelmente se reflectiu na redução da frequência de outros grupos taxonómicos, visto que as Cianobactérias são reconhecidas como principais fontes de azoto fixado nestes ecossistemas, e a existência e bio-disponibilidade deste elemento tem sido apontada como um factor preponderante na abundância de microrganismos.

Em paralelo, este estudo também teve como objectivo avaliar a distribuição e afinidades filogenéticas de um grupo de microorganismos em particular, com um papel relevante no ciclo do azoto. Partindo do facto do Vale Vitória ter sido recentemente amostrado, a existência de bactérias (AOB) e arqueias (AOA) oxidadoras de amoníaco foi avaliada numa vasta gama de ecossistemas terrestres da Antártida, baseado em técnicas tradicionais de análise de DNA (extracção, amplificação, clonagem, sequenciação). Os locais amostrados foram os vales Miers, Beacon, Upper Wright, Battleship Promontory, e a região do glaciar Darwin-Hatherton, perto das Montanhas Darwin. O gene funcional *amoA* que codifica a sub-unidade alfa da enzima ammonia-monooxigenase foi o marcador de estudo, e os resultados revelaram uma baixa diversidade genética, com apenas 6 e 5 OTUs identificadas para AOB e AOA respectivamente com o clustering cut-off mais alto, e apenas 2 e 3 OTUs identificadas com o clustering cut-off mais baixo. As OTUs de AOA estão afiliadas com o grupo terrestre 1.1b, ao passo que as OTUs de AOB se encontram separadas em dois grupos, estando um afiliado com o género *Nitrosomonas*, e o com o género *Nitrosospira*. Tanto para AOA como para AOB parece existir uma OTU cosmopolita, que é mais abundante e tem afiliações com clones ambientais provenientes de uma vasta gama de ecossistemas. Os nossos resultados demonstram uma distribuição dispersa de AOA e AOB ao longo das Montanhas Transantárticas e reforçam o potencial dos processos de nitrificação em regular a funcionalidade microbiana nestes ambientes extremos.

Palavras-chave: Antártida, Vales Secos, Diversidade microbiana, Diversidade funcional,16s rRNA, *amoA,* nitrificação, tecnologias de nova sequenciação.

Abstract

Ice free regions in Antarctica account for approximately 0.3% of the continental area and are patchily distributed. The extreme environmental conditions of these soil ecosystems have shaped a low diversity and simple trophic structure in which microorganisms face severe conditions including low water and nutrient availability, cold temperatures, freeze-thaw cycles, long periods of darkness in winter, and exposure to high levels of ultraviolet radiation in summer. In these areas, the diversity of microorganisms involved in key biogeochemical processes such as the nitrogen cycle is still largely unknown.

This study is integrated in a multidisciplinary research team based in New Zealand (NZTABS) that focuses primarily on the microbial diversity of the Antarctic Dry Valleys. Given this, the K020 expedition to Victoria Valley in January 2013 sampled an area of approximately 300 km^2. Soil samples were collected, along with the corresponding metadata and subjected to automated ribosomal intergenic spacer analysis (ARISA), to characterize the abundance and richness of the existing microbial communities in the whole valley. Relationships between environmental and microbial data from a total of 86 stations revealed that diversity, abundance and activity of the Victoria Valley's bacterial communities are subjected to strong spatial structuring due to the extreme heterogeneity in soil geochemical properties, being water availability identified as a main environmental constrain.

In order to understand how water availability in these environments may affect bacterial diversity and phylogeny, soils from a transect with increasing distance from a water source were also sampled and the 16S rRNA gene was amplified and sequenced using Roche's 454 pyrosequencing technology. Alpha and beta diversity metrics were computed and compared with measures of environmental variables which were also collected on site. Results revealed a clear shift in the microbial communities across the transect with a water availability gradient, namely a notorious replacement of members of the Proteobacteria phyla by members of the Actinobacteria phyla, as the soils became drier. Additionally, the sampling point with most unknown OTUs corresponded to the place with the highest values of conductivity and organic carbon production, along with the highest frequencies of members of phylum Bacteroidetes. Also, members of the Cyanobacteria phyla decreased their frequencies as the soils became drier, which likely caused a decrease in the frequencies of occurrence of other taxonomic groups, given the fact that Cyanobacteria are known to be the major nitrogen fixators in the Dry Valleys

and the bio-availability of this chemical element has proved to influence the abundance of other microorganisms.

In parallel, this study also aimed to assess the distribution and phylogenetic affinities of a particular group of microorganisms with a prominent role in the nitrogen cycle. Taking advantage of the recently sampled Victoria Valley, the existence of ammonia-oxidizing bacteria (AOB) and archaea (AOA) was assessed in a high range of Antarctic soil environments, based on standard DNA molecular approaches (extraction, amplification, cloning and sequencing). We established relationships between the broad scale environmental gradients of the sampled locations and the relative diversity of ammonia oxidizing microbial communities. In addition to Victoria Valley, the Miers Valley, Beacon Valley, Upper Wright Valley, Battleship Promontory, and the Darwin-Hatherton Glacier region of the Darwin Mountains were the chosen sampling sites, from which DNA was extracted, amplified, cloned and sequenced.

The gene coding for ammonia-monooxygenase (*amoA*) was the functional marker chosen and results revealed generally low AOB and AOA *amoA* gene diversity, with only 6 and 5 identified OTUs with the highest clustering cut-off, and, and 2 and 3 OTUs respectively with the lowest clustering cut-off. AOA OTUs were affiliated with the soil group 1.1b, along with culturable representatives of the genus *Nitrososphaera*. AOB OTUs were separated into two clusters, one affiliated with the genus *Nitrosomonas*, and another one affiliated with the genus *Nitrosospir*a. In both groups of microorganisms there seems to exist a cosmopolitan OTU, which is the most abundant and has close affiliations with environmental clones from a wide range of environments. The observations reported in this study demonstrating a highly dispersed distribution of AOA and AOB within the Transantarctic Mountains reinforce the potential of nitrification processes in driving microbial functionality of these extreme ecosystems.

Key words: Antarctica, Dry Valleys, Microbial diversity, Functional diversity, 16S rRNA, *amoA,* nitrification, next-generation sequencing technologies.

Table of contents

List of papers

The elaboration of this dissertation benefited from the following submitted scientific publications and additional conference presentations with a relevant contribution:

1) Monteiro, M., **Séneca, J**., Magalhães., C.: The history of aerobic ammonia oxidizers: from the first discoveries to today. *Journal of Microbiology* 2014, 52 (7), pp. 537-547

2) Magalhães, M., Machado, A. Monteiro, M. **Séneca, J.,** Charles, L., Cary, C.S.: Environmental constraints on the distribution of key nitrogen genes in the extreme Transantarctic Mountains. Submitted to *Frontiers in Microbiology*

3) **J. Séneca**, M. Monteiro, H. Ribeiro, L. Charles, S. C. Cary and C. Magalhães: "Phylogenetic diversity of ammonia-oxidizing microorganisms in the extreme Transantarctic Mountains" - 6ª Conferência Portuguesa de Ciências Polares - October 31st 2014 Porto, Portugal. (Oral)

4) **J. Séneca**, M. Monteiro, H. Ribeiro, L. Charles, S. C. Cary and C. Magalhães:"Antarctic Dry Valleys: geochemical soil properties and microbial communities" - *IJUP 2014* (Seventh Meeting of Junior Researchers of the University of Porto) - February 12th 2014, Porto Portugal. (Oral)

5) **J. Séneca**: "Diversity and abundance of Bacteria vs Archaea nitrifiers in the extreme Dry Valleys of the Transantarctic Mountains" - *IV Workshop APECS Portugal - How To Be a Polar Scientist for Dummies*. October, 31th 2013. Anfiteatro Verde, Faculdade de Ciências e Tecnologia, Piso 1, Edifício 8, Faro. (Oral)

6) **J. Séneca**, M. Monteiro, H. Ribeiro, L. Charles, S. C. Cary and C. Magalhães "NITROEXTREM: 2013 Antarctic Field Campaign to the Dry Valleys" - *5ª Conferência Portuguesa de Ciências Polares* - November 1st 2013 Faro, Portugal. (Poster)

List of figures

List of tables

List of abbreviations

®	Registered trademark
16S rRNA /SSU rRNA	16S ribosomal RNA / Small subunit of ribosomal RNA
AMO	Ammonia monooxygenase
amoA	Gene coding for the αsub-unit of the ammonia-monooxygenase
AOA	Ammonia-oxidizing Archaea
AOB	Ammonia-oxidizing Bacteria
ARISA	automated approach for ribosomal intergenic spacer analysis
ATP	adenosine triphosphate
A_w	Water activity
BLAST	Basic Local Alignment Tool
bp	base pair(s)
CTAB	bromide-polyvinylprrolidone-b-mercaptoethanol
DAPI	4',6-diamidino-2-phenylindole
DDBJ	DNA Databank of Japan
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
dNTP's	Deoxynucleotide Triphosphates
e.g.	*exempli gratia*
EMP	Earth Microbiome Project
ENA	European Nucleotide Archive
FU	Fluorescence Units
GG	GreenGenes
H'	Shannon-Wiener diversity index
HAO/*hao*	Hydroxilamine oxidoreductase
IC	Inorganic Carbon
ICTAR	International Centre of Terrestrial Antarctic Research
ITS	Internal Transcribed Spacer
mM	Milimolar
M.y.	Million year
N	Nitrogen
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing

nifH	Gene coding for nitrogenase
NJ	Neighbour-Joining
NZTABS	New Zealand Terrestrial Antarctic Biocomplexity Survey
ºC	Celsius degrees
OC	Organic Carbon
ºE	Degrees East
ºN	Degrees North
ºS	Degrees South
OSD	Ocean Sampling Day
OTU	Operational Taxonomic Unit
ºW	Degrees West
PAR	Photosynthetically active radiation
PcoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
RDP	Ribosomal Database Project
RNA	Ribonucleic acid
rpoB	Gene coding for the β-subunit of RNA polymerase
S	Species Richness
TC	Total Carbon
™	Trademark
T-RFLP	Terminal Restriction Length Fragment Polymorphism
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
USA	United States of America
UV	Ultra-violet radiation
v.	Version
vs	versus
WGS84	World Geodetic System 1984
µL	Microliter
RLU	Relative Light Unit
µS	Microsiemens

Introductory Review

1. Environmental Microbiology

Microorganisms are recognized as the most ubiquitous and widespread living group on Earth, inhabiting all ecosystems of our planet. They are known to play a major role on Earth's biogeochemical cycles, plant nutrition, symbiotic relationships and primary production. Multiple studies have come to the conclusion that the majority of known microorganisms cannot be cultured by standard techniques, and that this uncultured fraction (approximately 99%) included diverse organisms, distantly related to the cultured ones (Torsvik et al. 1990; Riesenfeld et al. 2004). Hence, the existing pure cultures are not representative of the existing biodiversity, and even though there has been a boost in the application of culture independent

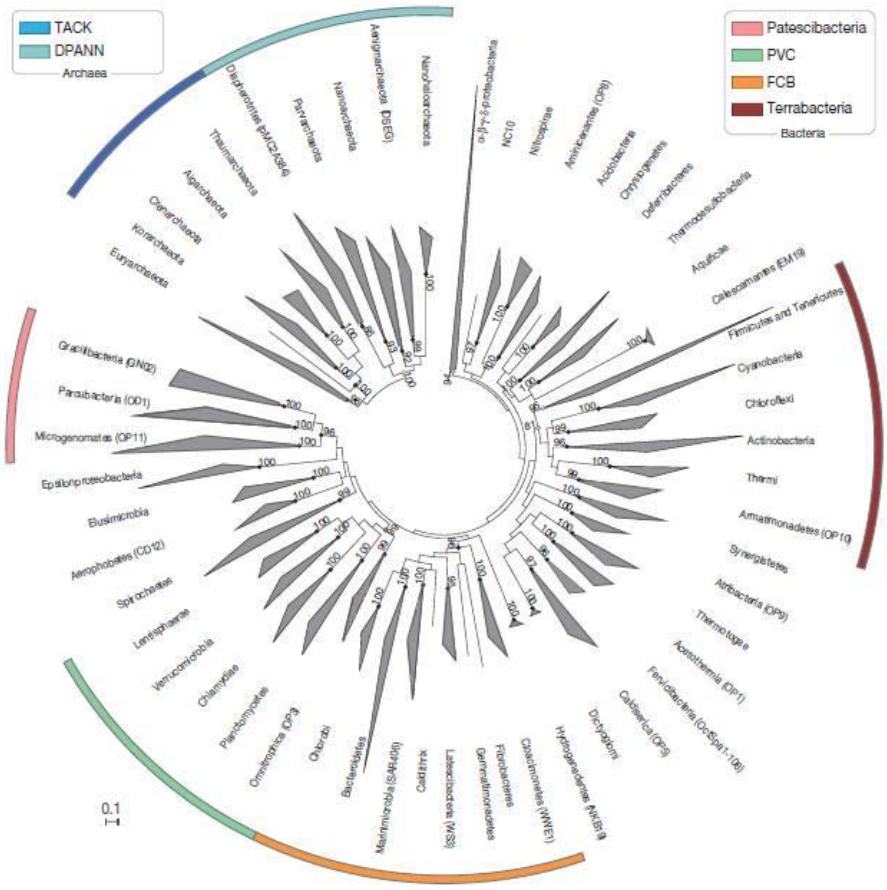


Figure 1. Maximum likelihood phylogenetic inference of the Bacteria and Archaea domains. This tree is based on up to 38 marker genes and collapsed at the phylum level. Superphyla (TACK, DPANN, FCB, PVC, Terrabacteria, and Patescibacteria) are highlighted with color ranges. Adapted from Rinke et al. (2013)

methods to the field of environmental microbiology, a high percentage of the genetic diversity, community composition, relative abundance, and distribution of microorganisms remain under sampled, and uncharacterized (Handelsman 2004; Riesenfeld et al. 2004; Sogin et al. 2006).

Naturally, culture independent studies boosted after the discovery of the DNA structure (Watson & Crick 1953), the genetic code (Nirenberg et al. 1966) and the quick development of pioneering sequencing techniques (Sanger & Coulson 1975; Sanger et al. 1977). By focusing on natural communities, the studies of Pace and colleagues represented a turning point in the way we see the microbial world (Olsen et al. 1986; Pace 1997). Not only there was an acknowledgement that the main diversity of life was microbial, but Pace and colleagues managed to distribute microbial diversity into three primary domains (Archaea, Bacteria and Eukarya), based in the extraordinary works of Woese and colleagues regarding the potential of the 16S rRNA gene to encode evolutionary signatures (Woese & Fox 1977; Woese et al. 1990; Pace 1997).

Regarding prokaryotes, recent studies have shown that there are at least 50 bacterial phyla, and half of them are composed entirely of uncultured bacteria (Schloss & Handelsman 2004). The same situation applies to Archaea, with five recognized phyla and large amounts of unclassified environmental sequences (Gribaldo & Brochier 2009; Brochier-Armanet et al. 2011) (Fig. 1). Hence, the quest for the prokaryotic tree of life is still running and represents a fundamental question in microbiology (Gribaldo & Brochier 2009). The tree of life allows a clearer understanding of the deepest events in the history of Life on Earth by pointing out the evolution of microbial diversity over geological time, and with this, the emergence of important metabolic capacities which shaped early ecosystems (Delsuc et al. 2005; Pace et al. 2012).

Consequently, the reconstruction of the prokaryotic tree of life enables the unfolding of the evolutionary history of particular cellular processes (metabolic pathways) or systems (macromolecular complexes), by analyzing their components and thus represents an important and promising branch in phylogenomics (Delsuc et al. 2005).

2. Soil Microbiology

The analysis of microbial populations in natural habitats is one of the cornerstones of current research regarding the functioning of natural ecosystems, and soils represent no exception. The high variability of the soil's physicochemical properties, namely age, depth, mineral and organic elements, enables it to harbor most of the still uncharacterized microbial diversity (Torsvik & Øvreås 2002; Mocali & Benedetti 2010; Fierer et al. 2012; Janssen & Prosser 2013). Additionally, there are no straight relationships between soil's heterogeneity, biochemistry, spatial and temporal variability, and the existing microbial communities, since they are randomly spread, following nutrient gradients, moisture contents and pH (among others), showing a "hot-spot" or patchy distribution that makes representative sampling more difficult (Nunan et al. 2002). Also, the complexity of the existing microbial communities and their biotic and abiotic interactions is not fully understood, and varies among ecosystems (Leininger et al. 2006; Jia & Conrad 2009; Magalhães et al. 2009; Daebeler et al. 2012; Ladau et al. 2013).

The prokaryotic community in soils, like in any other environment, is composed of two domains - Eubacteria and Archaea – which present different abundance and diversity levels, depending on local soil characteristics (Torsvik et al. 1990; Torsvik & Øvreås 2002). Because of this, it is not possible to precisely point out one prominent ubiquitous taxon common to every soil on Earth.

Prokaryotes in soils are also responsible for mediating important pathways in most of the Earth's major biogeochemical cycles such as the nitrogen (Nicol & Schleper 2006; Hayatsu et al. 2008), carbon (Baker et al. 2013), and the sulfur and phosphorus cycles (Falkowski et al. 2008). In addition, they are also responsible for maintaining soil aggregation (Chotte 2005). Moreover, most soil microorganisms in bulk soil are in a dormant state, but readily burst into activity when water or easily decomposable substrates become available, which is followed by a succession of microfaunal predators such as protozoa, promoting an interplay between rates of nutrient cycling and a strong enhancement on the availability of mineral nutrients to plants (Scheu et al. 2005).

Similarly to other environments, culture-dependent methods presented a major drawback in the estimation of microbial diversity and abundance, a phenomenon known as the Great Plate Count Anomaly (Hugenholtz 2002). Hence, and with advances in molecular techniques, microbiologists realized that the "snapshot" produced by the application of culture-independent methods provided a much more realistic framework of the molecular make-up of

whole complex soil communities, as well of specific microorganisms and genes therein (van Elsas & Boersma 2011)

3. Methods to assess soil microbial diversity

Microbial diversity describes complexity and variability at different levels of biological organization. It encompasses genetic variability within taxa (species), their number (richness), relative abundance (eveness) and functional groups (guilds) in communities (Torsvik & Øvreås 2002). The species concept is a recurrent controversial issue that is common to many biological disciplines. The prokaryotic species concept has been developed in parallel to the design of molecular techniques that allowed the retrieval of more accurate information than the one previously retrieved from morphological and physiological data. However, taxonomists and microbiologists have not reached a consensus relatively to what is considered a prokaryotic species, since traits as different as ways of obtaining energy (heterotrophy *vs* autotrophy, among others), trophic roles, and different DNA sequence identity clustering thresholds have to be considered (Konstantinidis et al. 2006; Fierer et al. 2007; Schloss & Westcott 2011). Nevertheless, the present working hypothesis for the prokaryotic species concept is that microorganisms belonging to the same species are a cluster of strains which are more related in terms of sequence identity and gene content, among themselves, than to strains outside the cluster (Konstantinidis et al. 2006; Chan et al. 2012; Cordero & Polz 2014). In addition to the species concept, the prokaryotic “ecotype” concept is defined as a strain or a group of strains that show some level of ecological distinctiveness, even though they belong to the same species (Konstantinidis et al. 2006; Rocha 2008).

The vast majority of current molecular analysis from soil is preceded by direct DNA/RNA extractions (Zuckerkandl & Pauling 1965; Torsvik et al. 1990). Key issues regarding this method are that it is highly reproducible and provides easy access to genes of extant soil microbial communities. However, the chemical integrity and purity of soil DNA can bias and/or limit further downstream analysis such as polymerase chain reaction (PCR).

A major step forwards regarding the study of soil microbiota via DNA (and/or RNA) has been the development of direct PCR amplification of target genes, such as the 16S rRNA (Wilson et al. 1990) and *rpoB* (Dahllöf et al. 2000) genes, and functional gene markers such as *amoA* and *nifH* (Rotthauwe et al. 1997; Jung et al. 2011). Despite being the molecular technique of choice, PCR amplification of soil DNA can be hampered by soil intrinsic enzymatic inhibitors. In addition, the perceived diversity is prone to differential amplification, meaning that particular targets amplify at higher rates than other, biasing results against the

so called “rare biosphere” (van Elsas & Boersma 2011). This can be overcome by using group specific primers that target low-abundance microorganisms. Combined with nested or semi-nested techniques, this approach helps to provide greater insights regarding the ecology of the target groups, as they reduce the complexity of the target community (van Elsas & Boersma 2011; Burke et al. 2011)

When working with environmental samples, community fingerprinting techniques such as denaturing gradient gel elecrectrophoresis (DGGE), terminal restriction fragment length polymorphism (T-RFLP), and automated approach for ribosomal intergenic spacer analysis (ARISA) work as a proxy for bacterial community composition and are useful in order to have a preliminary idea of microbial sample composition (Smalla et al. 2007). DGGE is a technique that was implemented in the early nineties, and is based on the mobility of nucleic acids in a denaturing gel of acrylamide. It relies on the profiles of double stranded PCR products (Muyzer et al. 1993). Hence, sequences are separated according to their GC % (Muyzer et al. 1993). T-RFLP (Liu & Marsh 1997) and ARISA (Fisher & Triplett 1999) are community fingerprinting techniques based on the PCR amplification of target genes using fluorescently labeled primers. The resulting output is an electropherogram showing a series of peaks relating fragment length with fluorescence intensity (which reflect the abundance of certain taxa), and

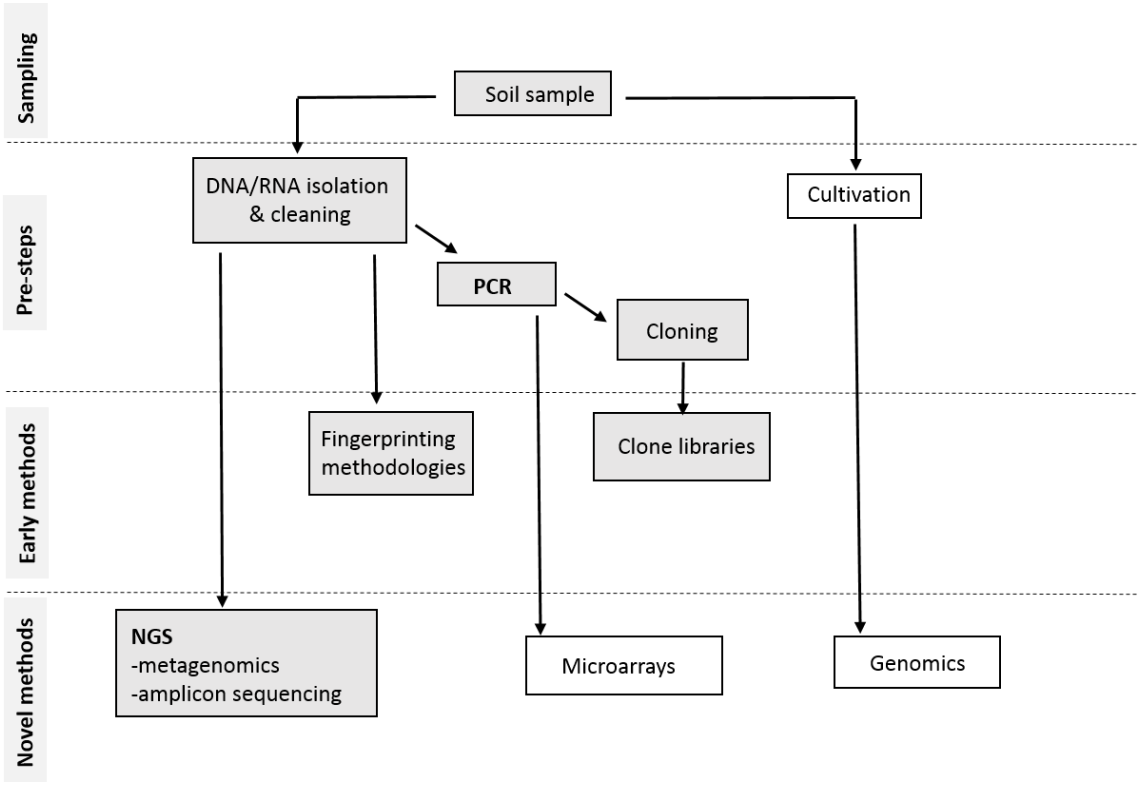


Figure 2. Schematic representation of a general workflow to assess microbial diversity in soils. Grey boxes represent steps used in this project. Modified from Elsas and Boersma (2011).

from them, with the appropriate statistical tools, several diversity indexes can be calculated (Boon et al. 2002; Culman et al. 2009).

The rRNA genes have, since the end of the 70's, become the molecular marker of choice, when it comes to infer phylogenetic relationships at different levels, among different taxonomic levels (Zuckerkandl & Pauling 1965; Ward et al. 1990; Weisburg et al. 1991). They are greatly involved in the protein-synthetizing process within the cell, and tend to show a high level of conservation among all organisms, regarding structure, functionality, and nucleotide sequence. In addition, there is no evidence of horizontal gene transfer between contemporaneous organisms, and therefore, relationships between rRNAs reflect evolutionary relationships between the organisms (Olsen et al. 1986). Regarding prokaryotes, the 5S rRNA and the 16S rRNA genes have been used to assess diversity at different levels (Olsen et al. 1986; Riesenfeld et al. 2004). However, given the fact that the 16S rRNA gene sequence is longer, this marker is preferable over the 5S rRNA gene when it comes to inferring phylogenetic relationships, especially with the development of DNA cloning/sequencing techniques that enable the use of the total 16S gene sequence, instead of partial sequences (Olsen et al. 1986). Nevertheless, it has some pitfalls and among them is the fact that it exists in more than one copy in certain microbial taxa, which leads to an overestimation of the existing microorganisms (Vos et al. 2012; Poretsky et al. 2014). Therefore, 16S RNA-based community analysis should be carefully performed and the use of single copy genes such as *rpoB* should be used instead or as a complementary tool (Farrelly 1995; Dahllöf et al. 2000; Case et al. 2007). This approach, however, is still limited since the amount of 16S RNA gene sequences in public repositories is huge, when compared with any other gene (Maidak et al. 1996; Cole & Tiedje 2014).

Regardless of the potential of the 16S rRNA gene, its use as a standalone fingerprint tool has been progressively replaced (Riesenfeld et al. 2004). The use of high throughput shotgun sequencing of environmental samples – metagenomics – has enabled the recovery of massive amounts of genetic data present in a given habitat, providing clues regarding the functional capacity of a given community, rather than just its phylogenetic composition. When compared with the use of the 16S rRNA gene, the analysis of full genomes by metagenomic approaches enables researchers to go one step forward towards the understanding of microbial functionality in different ecosystems (Walker et al. 2010; Fierer et al. 2012; Lebedeva et al. 2013). 16S rRNA gene profiling using next-generation sequencing (NGS) techniques is often used as a preliminary step before metagenomic analysis and is useful to determine the taxonomic composition of the samples, and can also be useful in guiding decisions regarding

a preferential sequencing technology (Roche 454/Illumina/Ion Torrent), and the amount of sequencing necessary (Klindworth et al. 2013; Poretsky et al. 2014; Mori et al. 2014).

Roche 454 was the first commercially successful next generation system and uses pyrosequencing technology, that is, it relies on the detection of pyrophosphatase released during nucleotide incorporation, instead of using traditional dideoxynucleotides (ddNTPs). DNA with 454-specific adapters is denatured into single strand and captured by amplification beads that precede emulsion PCR (Parameswaran et al. 2007). Then, in a picotiter plate, the dNTPs are sequentially incorporated and different amounts of pyrophosphate are released, resulting from the incorporation of different nucleotides. Visible light is produced, registered, and a pyrogram profile is generated (Liu et al. 2012).

Other sequencing technologies exist, namely Illumina and Ion Torrent that use different sequencing chemistry settings and workflows (Parameswaran et al. 2007; Liu et al. 2012). Additionally, single-cell metagenomics has received considerable attention in recent years, being considered the 2013 method of the year by the Nature Publishing Group (Nature Methods, 2014). This technique makes rare cells more accessible to analysis, provided that methods are available to isolate and/or enrich these cells from their heterogeneous environments, which is still a major challenge.

The use of sequencing technologies is evolving at an unprecedented rate. A decade ago, the sequencing cost of a raw megabase of DNA rounded $1K, and that cost has now decreased more than 10 fold (Segata et al. 2013; Wetterstrand 2014) which has resulted in more and more researchers using metagenomics as a potential answer to their biological questions. However, other areas must also follow up these advances, namely Bioinformatics. As an example, while the amount of data/sequences generated has increased 5000 fold, the computational speed has only increased 10-fold, which results in a high increase in the time needed to analyze full datasets (Stein 2010). Additionally, means to store data (and associated metadata) must also be generated and be of easy access. Large sequence repositories already exist and are expanding, namely ENA (European Nucleotide Archive) (Leinonen et al. 2011), NCBI (National Center for Biotechnology Information) (Geer et al. 2010), and DDBJ (DNA Databank of Japan) (Tateno & Fukami-Kobayashi 1998). These databases are expected to store data from ambitious massive worldwide sampling surveys such as the Earth Microbiome Project (EMP) and the Ocean Sampling Day (OSD) and to conveniently make it available to multidisciplinary teams throughout the globe, in order to produce a global Gene Atlas describing microbial diversity, protein and environmental models for each biome (Gilbert et al. 2014; www.microb3.eu/osd)

In conclusion, the advent of new sequencing methods has proved to enable a paradigm shift in the field of microbial ecology, unravelling crucial information regarding the genetic features and ecosystem role of countless microorganisms in all of the Earth's microbiomes. The dual realization that there are still so many unknown microorganisms out there and that researchers now have the tools to discover and characterize them represents a turning point in microbiology studies, opening new doors of exciting research.

4. Microbial life at the extremes

4.1 Cold adapted organisms

The term "extreme" is relative and is always compared to what is considered "normal". An extreme environmental condition is defined as a condition that goes beyond acceptable ranges that have been observed throughout time, in a given region/ecosystem. Extreme environments include high gradients of temperature, pH, nutrient and water availability, levels of radiation, heavy metals, and toxic compounds (Satyanarayana et al. 2005). Organisms that can thrive in these harsh conditions are named "extremophiles" and the majority of them are included in the microbial world (Morita 1975; Schafer 2004)

About 85% of the biosphere is permanently exposed to temperatures below 5 ºC. These habitats extend from the Arctic to the Antarctic and their major fraction is represented by deep sea, snow, permafrost, sea ice and glaciers (Margesin & Miteva 2011).

Psycrophiles are true extremophiles, as they are mostly adapted to low temperatures (have an optimum growth temperature < 15ºC and can grow at or below 0ºC) but are also found in environments with other constrains, as is the case of microorganisms that inhabit deep ocean environments, in which pressure levels are very high (Morita 1975; Feller & Gerday 2003). In order to thrive in such harsh conditions, psychrophilic organisms have to overcome some physiologic problems related with the environment they live in, namely at the cell membrane level. At low temperatures, enzymes become rigid and solute concentrations may increase to toxic levels. Additionally, once the water freezes, ice crystals may pierce the membrane. Hence, psychrophilic membranes contain higher levels of unsaturated lipids, when compared with their mesophilic counterparts, in order to balance membrane fluidity (Feller & Gerday 2003; Rampelotto 2010). Psychrophiles also produce cold-adapted enzymes that are characterized by a lower core hydrophobicity, fewer ionic and electrostatic interactions, a change in surface residues that increase solvent interactions, additional surface loops,

modification in residue composition (particularly a substitution of proline by glycine, and smaller arginine/lysine ratios), fewer interdomain/subunit interactions and aromatic interactions (Deming 2002; Rampelotto 2010). In addition, anti-freeze proteins have been identified, preventing cell membrane piercing by ice crystals (Feller & Gerday 2003). These factors contribute to a higher thermo stability necessary to withstand the existing low temperatures.

Major biotechnological advances are emerging from the study of physchrophiles, with direct applications at the industrial, health and environmental levels (Rothschild & Mancinelli 2001; Cavicchioli & Siddiqui 2002). Also, regarding the fields of biology and evolution, they provide a great ecological framework to address issues such as the origin of life, and the molecular mechanisms of cold adaptation (Cowan, 1997).

4.2 Antarctica

Antarctica was the last continent to be discovered, and is thought to have been spotted for the first time in 1820 by Nathaniel Palmer, though earlier descriptions that date to Ptolomy (1st century AD) had been made of a southern cold land – Terra Australis. Etymologically, the word Antarctica comes from the greek “antarktiké”, that means “opposite to the Arctic” and was formally attributed to the continent around 1890 by the Scottish cartographer John George Bartholomew.

It is often viewed as one of the Earth’s last great frontiers; the combination of extreme conditions with periods of complete or near darkness provides inhospitable conditions to the existence of life (Convey et al. 2008). Being the largest reservoir of fresh water on Earth (90%), with an area of 13.6 million km^2, a mean elevation of 3000 m above sea level, average winds of 80 km/h, an annual precipitation of less than 5 cm, and temperature ranges from -40 to -50 ºC, Antarctica is a continent of extremes (Bockheim 1997; Cowan 2014). The existence of such extremes means that the influence of the Antarctic region on both climate and oceans extends not only to its immediate area, but also into mid-latitude global systems. In fact, Antarctica plays a remarkable role in the global environmental system in terms of climate, global heat balance, oceanic circulation and marine nutrient cycling. Spatial and temporal variation in these systems is crucially important not only to human understanding of how the planet currently functions, but also to predict future changes (Doran et al. 2002).

Historically, Antarctica emerged from the disintegration of the Gondwana supercontinent in the Cretaceous (~ 120 M.y. ago) and became fully isolated around ~33.7 M.y., when it separated from South America (Clarke et al. 2005). Understanding the geology of Antarctica is hampered by the lack of rocky outcrops, which are thought to account for less

than 0.3 % of the continental area (Cowan 2014). However, nowadays three regions can easily be distinguished in continental Antarctica: East Antarctica is generally above sea level, with a very thick ice-sheet cover. It has an active volcano – Mount Erebus – and one of the world's biggest lakes – Lake Vostok – whose waters have been isolated from air and light, under the pressure of the ice-sheet cover, therefore providing a good setting for the discovery of new forms of microbial life, as it has recently been proved (Shtarkman et al. 2013). East Antarctica is a stable, ancient shield of Precambrian metamorphic and igneous rocks, formed mostly between 3800 and 500 M.y ago (Harley 1988). Younger folded sediments of Upper Proterozoic-Lower Palaezoic age are exposed in the fold belts that border the Precambrian metamorphic shield (Talarico & Kleinschmidt 2008). West Antarctica is lower than the East part, comprising areas that are below sea level. It also presents a more heterogeneous topography and is dominated by Mesozoic and Cenozoic rocks. Separating these two parts, there are the Transantarctic Mountains, which are one of the Earth's major mountains. They reach 4000 meters in height, 100-200 km in width and extend for some 3500 km across the continent, between the Ross and Weddell seas (Cary et al. 2010).

Due to its remoteness and glaciation, Antarctica is one of the Earth's most pristine environments, and provides essential data for the construction of a global baseline against which to monitor climate-related changes that may affect global processes. Particularly regarding recent climate changes, Antarctica has been pointed as the continent where these changes will be (and already are) first felt (Doran et al. 2002; Talarico & Kleinschmidt 2008) and therefore there is an increasing need to understand the dynamics of Antarctic ecosystems to predict future climate change scenarios and to properly account for their conservation (Seybold et al. 2009; di Prisco et al. 2012).

4.3 Ice-free areas in continental Antarctica

Approximately 99.7% of continental Antarctica is permanently covered by the Antarctic Ice sheet throughout the year and the few remaining areas that are ice-free are patchily distributed across the continent (Fig. 3) (Hopkins et al. 2006). Antarctic terrestrial ecosystems are different from other polar ecosystems (namely the Arctic), in the sense that they are colder, drier, lower in nutrient availability, and often alkaline, since soils accumulate salts from rare precipitation events and weathering, due to extreme aridity (Cary et al. 2010; Margesin & Miteva 2011). These environmental conditions have shaped soil ecosystems of low diversity and simple trophic structure (Pointing et al. 2009; Fierer et al. 2012). With the exception of the Antarctic Peninsula and some sub-Antarctic islands, biodiversity is mostly restricted to invertebrates, protozoa, fungi, Bacteria and Archaea. Hence, Antarctic soil communities are mostly microbial and their structure seems mainly influenced by abiotic factors, due to the few existing biotic interactions (Hogg et al. 2006).

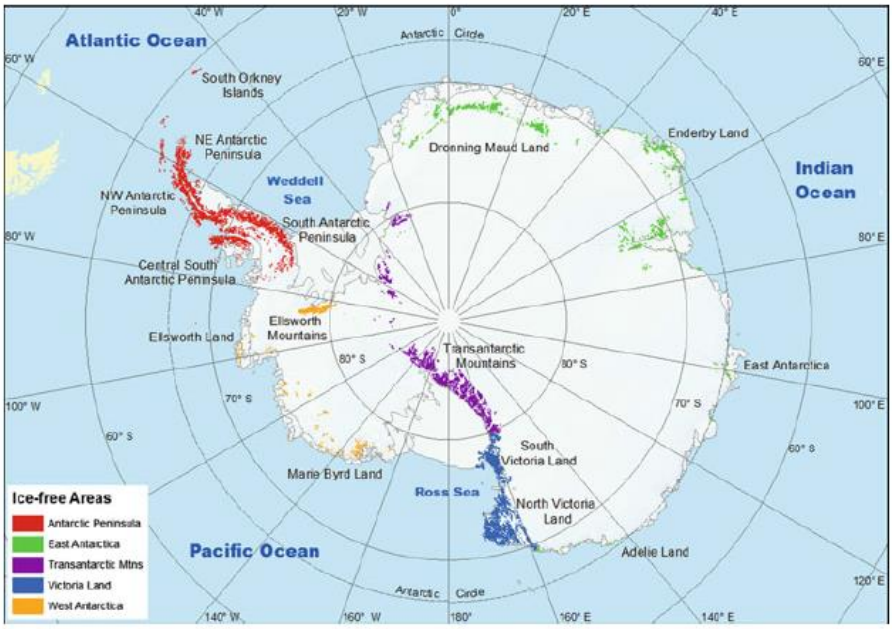


Figure 3. Ice-free regions in the Antarctic continent: Ice-free regions account for approximately 0.3% of the whole continent (Bottos et al. 2014)

4.4 The McMurdo Dry Valleys: the coldest and driest deserts on Earth

The Dry Valleys of Antarctica are a mosaic of glacier, ice-covered lakes, ephemeral streams and arid soils (Wall 2005; Barrett et al. 2006a). They are a series of generally west-to-east oriented, glacially curved valleys located between the Polar Plateau and the Ross Sea, in Southern Victoria Land. They are primarily ice-free because the glacial flow from the East Antarctic Ice Sheet is obstructed by the Transantarctic Mountains. The Dry Valleys are characterized by even steeper environmental gradients than continental Antarctica. They have an average mean temperature that ranges from -15 to -30. Arid soils are, however, subjected to large temperature fluctuations, namely during the austral summer, where temperatures can go up to 0 ºC (and liquid water is potentially available) and largely decrease at the end of the day. Temperature fluctuations of > 20 ºC are not uncommon, leading to multiple freeze-thaw cycles in the same day. In winter, minimum average temperatures range between -40 and -60 ºC (Zeglin et al. 2009; Cary et al. 2010).

Precipitation levels in the Dry Valleys are very low, presenting levels of < 10 cm^2/year and are mostly in the form of snow, which never reaches the soil due to rapid sublimation. This quick evaporation is driven by the very low relative humidity of the katabatic winds that come from the Polar Plateau (Chwedorzewska 2009). Additionally, mineral soils are also very dry, with mass water content below 2%, which is equivalent to the water content of many hot deserts. This effect is synergistically affected by the existing high levels of salinity that are the result of a continued aerosol deposition and very low leaching rates (Niederberger et al. 2008). In addition to the aforementioned constraints, factors such as low organic carbon content (from 0.0 to 0.43%) high incidence of radiation (namely UVB) and photosynthetically active radiation (PAR), absence of soil structure and cohesion, soil age and other geochemical properties also impose strong limitations on the survival, prevalence and distribution of soil biota (Sokol et al. 2013).

The trophic simplicity of these extreme environments provides a particularly good framework to infer relationships not only between intrinsic soil features, but also regarding the role of environmental parameters in shaping the distribution of the microbial communities (Mocali & Benedetti 2010).

4.5 Microbial diversity studies in the Dry Valleys

Early cultivation-based methodologies have successfully isolated and described many bacterial strains in Antarctica (Darling & Siple 1940; Flint & Stout 1960). However, as mentioned previously, these methods highly underrepresented the existing microbial diversity (Smith et al. 2006). As a result, Antarctic soils were initially classified as "hostile" and "unable to sustain life". The application of culture-independent methods (Brambilla et al. 2001; Wall et al. 2006; Cowan 2009; Cary et al. 2010) proved that Antarctic mineral soils, and particularly the Dry Valleys, harbor a considerable amount of microbial diversity, when compared to what was initially thought and to what was observed in other environments (Cary et al. 2010) (Fig. 4). However, even though these cold desert ecosystem harbor a much higher diversity than originally expected, microbial diversity is far lower in average than the one found in other biome types (Fierer et al. 2012).

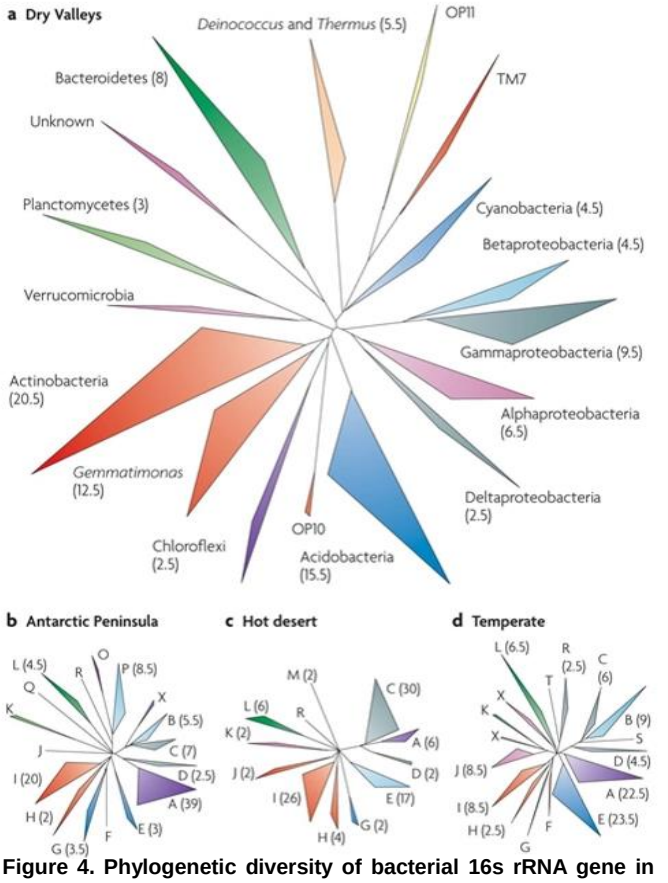


Figure 4. Phylogenetic diversity of bacterial 16s rRNA gene in different ecosystems. a. McMurdo Dry Valleys ((Smith et al. 2006; Niederberger et al. 2008; Aislabie et al. 2006); **b.** Antarctic Peninsula (Yergeau et al. 2007b); **c.** Hot desert (Chanal et al. 2006); **d.**Temperate surface soil (Liles et al. 2003; Lipson & Schmidt 2004; Sun et al. 2004). A, alphaproteobacteria; B, betaproteobacteria; C, gammaproteobacteria; D, deltaproteobacteria; E, Acidobacteria;F, OP10; G, Chloroflexi; H, Gemmatimonas; I, Actinobacteria; J, Verrucomicrobia;K, Planctomycetes; L, Bacteroidetes; M, Deinococcus and Thermus; O, TM7; P, cyanobacteria;Q, Fusobacteria; R, Firmicutes; S, epsilonproteobacteria; T, WS3; X, unknown isolates. (adopted from Cary et al. 2010)

A number of detailed phylogenetic surveys of Antarctic soils have been published recently (Aislabie et al. 2006; Smith et al. 2006; Yergeau et al. 2007b; Wood et al. 2008; Babalola et al. 2009; Pointing et al. 2009; Lee et al. 2012; Richter et al. 2014) and the most striking and consistent observation is that a substantial proportion of retrieved 16S rRNA gene sequences did not show high homology to the sequences of cultivable organisms, suggesting

that Antarctic soils harbor a large number of novel and possibly endemic species (Cowan 2009; Pointing et al. 2009; Cary et al. 2010; Sokol et al. 2013).

Bacterial community structure in the Dry Valleys varies considerably from the one observed in other regions of the continent (Fig. 4). At the phylum level, there is a decrease in the relative abundance of Proteobacteria and an increase in relative abundance of the Bacteroidetes and Actinobacteria phyla, when compared with the Antarctic Peninsula (Cowan 2014). Also, the prevalence of highly resistant members of the Deinococcus-Thermus and Gemmatimonadetes phyla in the mineral soils of the Dry Valleys distinguishes these soils from those with more temperate influences. However, even between valleys it is possible to observe structural differences between bacterial communities (Lee et al. 2012). Regarding Archaea, their existence has only been assessed in coastal soils near the Ross Sea (Aislabie et al. 2006; Ayton et al. 2010) using the 16S rRNA gene as molecular marker. A recent survey focused in the influence of soil properties in the diversity and distribution of archaeal diversity in the Dry Valleys revealed low overall richness, with most of the operational taxonomic units (OTUs) affiliated with the Thaumarchaeota phylum, and the remaining affiliated with the Euryarchaeota phylum (Richter et al. 2014).

These observations, focused on the major groups of prokaryotes came to reinforce the role of abiotic factors in driving microbial diversity, their role in Antarctic biogeochemical cycles, and also to guide future conservation planning strategies (di Prisco et al. 2012).

Research Objectives and Thesis Organization

Based on the premise that environmental variables have a preponderant role in shaping the abundance, diversity, structure and distribution of microbial communities in the Dry Valleys of Antarctica, this master project has two major goals which will be presented in two separate chapters:

Chapter I. The influence of the high gradients of soil geochemical parameters in controlling microbial metabolic activity, abundance and diversity was evaluated in Victoria Valley, covering a total of 86 sampling stations using community fingerprinting methods (ARISA). In order to understand in a more detailed way how one of those environmental parameters (water availability) dictates bacterial diversity and abundance, soils from a transect with increasing distance from a water source in Victoria Valley were retrieved and subjected to massive parallel pyrosequencing, by using Roche's 454 technology.

Chapter II. This chapter will present an inter-valley comparison within the Transantarctic Mountains, of the diversity and phylogenetic affinities of the bacterial and archaeal functional gene *amoA* which is involved in one of the Earth's major biogeochemical cycles – the nitrogen cycle. Here we investigated *amoA* gene diversity in the Darwin Mountains, Battleship Promontory, Miers, Beacon, Upper Wright and Victoria Valleys, where general soil microbial diversity and geochemistry have been previously described. We hypothesized that physicochemical heterogeneities of these extreme environments exert selective pressures on the groups of microorganisms involved in the ammonia oxidation pathway. Nitrogen biogeochemistry studies in the Antarctic Dry Valleys are still in their infancy and nitrogen is recognized to be a crucial element, regulating the microbial diversity and dispersion in these systems.

Chapter 1: *The influence of water availability in bacterial community structure: Victoria Valley as a case study*

I. Background

1. Victoria Valley: a geologic perspective

The Victoria Valley system is the largest of the McMurdo Dry Valleys, with an ice-free area of approximately 650 km^2. It contains five smaller valleys: the Victoria Upper, the Victoria Lower, Barwick, Balham and McKelvin valleys which include the Bull Pass area. These valleys are located between 400-1000 meters above sea level (Bockheim & McLeod 2013) (Fig. 5).This system of valleys is characterized by an internal drainage system, which channels seasonal meltwaters to its topographic low point, which is the 5.7 km^2 Lake Vida (McGowan et al. 2014).Three large glacier tongues flow from the valleys: Victoria Upper, into Victoria Upper Valley and the Webb Glacier into Barwick Valley from northwest; the Victoria Lower Glacier from the Wilson Piedmont Glacier and local alpine glaciers from the northeast; and a smaller alpine glacier (Packard Glacier), that flows into Victoria Lower Valley (Fig. 5).

There have been some scientific controversies regarding the topology of this valley system, namely the potential past extension and depth of Lake Vida, and the existence of other high level lakes (Calkin 1971; Hall et al. 2002). However, recent studies failed to validate this hypothesis, due to the lack of significant differences in the salt levels of soils in relation to elevation (Bockheim & McLeod 2013).

It is important from the ecological point of view to understand the geological features of these locations, as well as the chronology of their genesis, because the few existing life forms thrive mostly in soil, between and/or below the rocks, due to the low water availability and high incidence of ultraviolet radiation (Cameron 1972).

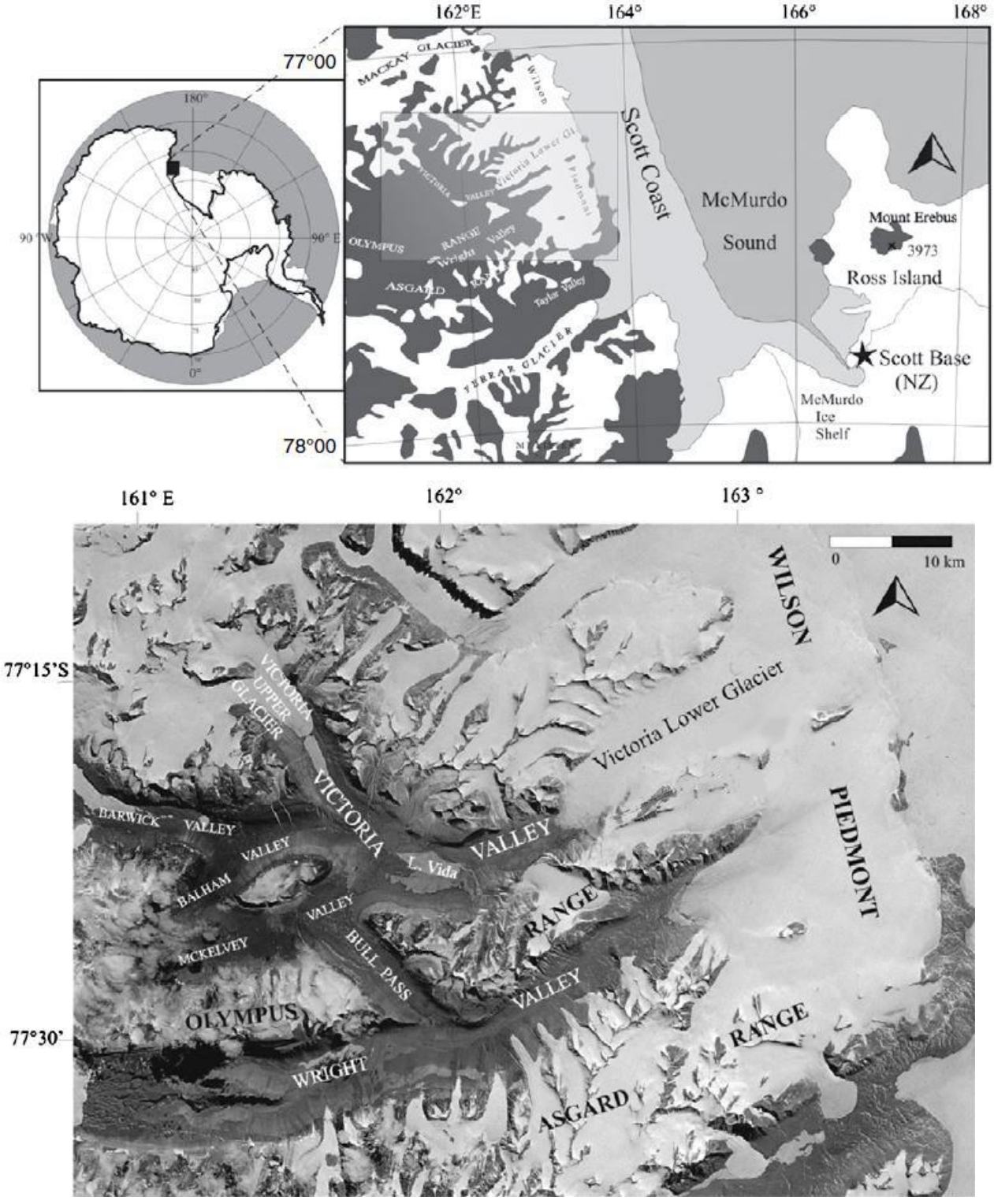


Figure 5. Location of the Victoria Valley System, Antarctica, showing the site of the bench-like features in the eastern Victoria Valley. The five valleys that compose the system are shown, as well as the biggest water source – Lake Vida – in the midpoint of the valley system. Adapted from McGowan et al. (2012).

II. Goals

Environmental parameters are known to shape the abundance and diversity of microbial communities. Using community fingerprinting methodologies (ARISA), we aimed to characterize the existing microbial communities in terms of bacterial richness and abundance in the Victoria Valley, from the 86 sampling stations sampled in the K020 Antarctic field campaign. In addition, given the fact that water is one of the major limiting factors to the existence of life in Antarctic ecosystems, we determined its effect in the bacterial community dynamics, in terms of taxa richness, abundance and diversity across a transect with a gradient of water availability in Victoria Valley. Inferences were done based on the 16s rRNA gene sequences amplified using a NGS technology (Roche's 454). NGS data analysis was performed using QIIME, but results from other two pipelines (SILVAngs and Metabiodiverse) are also presented.

We believe that this study provides critical insights regarding the intrinsically adapted and still largely unknown bacterial communities which are known to inhabit the Dry Valleys, which may represent sources of commercial, biological and medical interest.

III. Material and Methods

1. Sampling

The sampling program of this work was performed on behalf of the NITROEXTREM project, integrated in the ICTAR international program (ICTAR–www.ictar.aq), which coordinated, in January 2013, an Antarctic field campaign (K020) covering a sampling area of approximately 300 km^2 in the Victoria Valley system. One main field camp and three sub camps were set up to cover a total of 86 random sampling locations during three weeks of field campaign, with the main goal of building a model to link biodiversity with landscape and environmental factors. Soil from 86 sampling stations were collected (Fig. 6), and soil characteristics were measured for further analysis. In addition to the 86 sampling points, a total of six sites between T1 (77º 20.241'S, 161º 38.593'E (WGS84)) and T6 (77º 20.232'S, 161º 38.526'E (WGS84)) were sampled from a 32 m transect with increasing distance from a water pond near the main water source in this valley – Lake Vida. Several scoops of soil were collected aseptically and stored in a sterile Whirl-Pak (Nasco International Inc., Fort Atkinson, WI, USA) and kept at -80 ºC until further analysis.

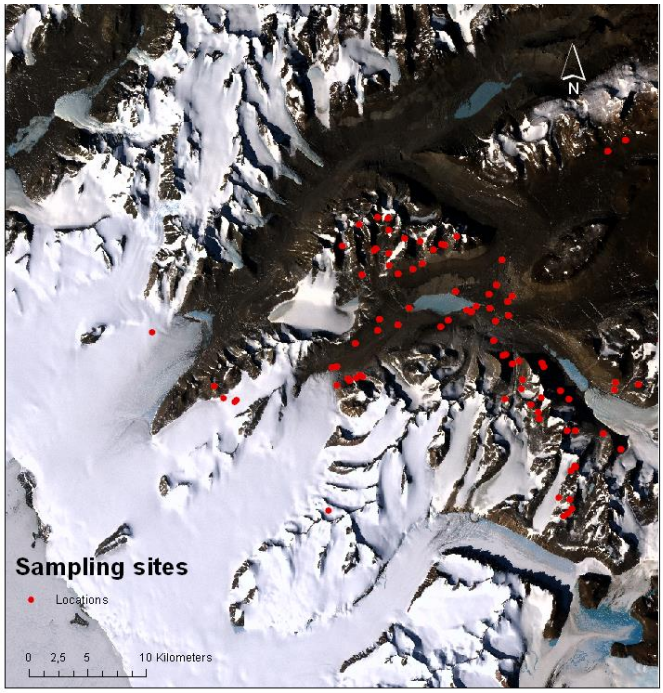


Figure 6.Location of the 86 sampling sites from which soil samples were collected in the K020 Antarctic field campaign.

All the procedures involved in the sampling campaign, as well as all the *in situ* measurements were performed and managed by the NZTABS in-field multidisciplinary team. All necessary and appropriate precautions were taken in order to avoid anthropogenic or cross-site contaminations. Aliquots of soil samples were shipped to Oporto's University and analyzed on behalf of this project.

2. Physico-chemical parameters

Water Activity (A_W) is a measure of the amount of water available to microorganisms and it is a very good indicator on whether the specific habitat is capable of supporting life. A_w in all sampling points was measured *in situ* using a portable water activity analyzer (Pa_WKit AquaLab, Decagon).

Conductivity and pH in all sampling points were also measured *in situ* using a CyberScan PC 510 Bench Meter (Eutech Intruments, Pte Ltd, Singapore) following the slurry technique which consists in mixing 1:2.5 mass ratio of samples and de-ionized water (Edmeades et al. 1985).

Organic carbon (OC) was only measured in the transect samples and was indirectly assessed from the subtraction of inorganic carbon (IC) from total carbon (TC). These two components were measured with a Shimadzu TOC-V CSH/CSN equipment. TC was measured by high temperature catalytic oxidation followed by non-dispersive infrared detection of CO_2. For IC, samples were acidified (1.5% HCl 2 M) and sparged with carrier gas (purified air) to convert only the inorganic carbon to CO_2, which was also detected by infrared. Calibration curves were generated for TC, using glucose (40% carbon) and for IC, using sodium carbonate (11.3% carbon). Three replicates from each sampling site were weighted (100 mg) and dried at 28 ºC for one hour before measurements.

3. Biological parameters

ATP (adenosine triphosphate) is a molecule that all living organisms produce and use in their metabolic processes, so the amount of biomass in a sample can be estimated by measuring the amount of ATP present in a sample. ATP was measured *in situ* by mixing the sample with an enzyme called luciferase, which reacts with the ATP to produce light. The amount of light produced indicates how much ATP is present in the sample and is measured in RLU (relative light units). In all the sampling points, this was done using a 3M™ Clean-Trace™ surface ATP kit, and results were measured in a 3M™ Clean-Trace™ NG Luminometer.

The total number of prokaryotic cells in the sediment samples was estimated by epifluorescence microscopy by a direct count of DAPI (4',6-diamidino-2-phenylindole - VWR) stained cells (Pernthaler et al. 2001; Llobet-Brossa, et al. 1998). One gram of sediment was weighed and fixed in 4% formaldehyde (AppliChem GmbH). Tween 12.5% (AMRESCO) was added and the sample was vortexed, sonicated (VWR® symphony™ Ultrasonic Cleaners) and

left for overnight incubation. In the following day, an aliquot of 350 µl of each sample was mixed with Tween 12.5 %, dilution water and 10 µl DAPI (0.5 mg/ml) (AppliChem GmbH) per ml of solution. Then, cells were filtered through the surface of 0.2 µm pore-size polycarbonate membranes by vacuum filtration and set up in a paraffin plate, to be later visualized using fluorescence microscopy (Zeiss Axiovision Z1). Fifteen fields were randomly photographed and cells were counted using the ImageJ 1.48v software (Abràmoff 2004). The following formula was used to calculate the total number of prokaryotes in the samples:

$$\text{Total cell counts} = \frac{a.b}{c.d} \times 10^6$$

Where,

a – area of the filter (mm^2)

b – average number of cells per field

c – field area (μm^2)

d – filtered volume (ml)

10^6 – conversion from mm^2 to μm^2

4. DNA extraction

Samples were extracted using a modification of the CTAB (Cetyltrimethylammonium bromide) extraction protocol (Barrett et al. 2006a). One gram of homogenized soil from each sample site was weighed and phosphate (100 mM NaH_2PO_4) and SDS buffers were added (100 mM NaCl, 500 mM Tris, pH 8.0, 10% SDS). After agitation, vortexing and recovering of the supernatant, the CTAB solution was added and samples were incubated for 30 minutes. Afterwards, they were washed twice with chloroform isoamyl alcohol (24:1), centrifuged and the supernatant was retrieved. Then, 7M of ammonium acetate buffer was added, samples were centrifuged, the supernatant was retrieved and 0.54 volumes of isopropyl alcohol were added. After centrifugation, and recovery of the supernatant, the samples were submitted to an overnight incubation. In the following day, samples were washed in 70% ethanol, eluted in 25 µl of nuclease-free water and stored at -80 ºC. Two replicates were extracted for each sampling site.

5. ARISA analysis

ARISA was used as a community fingerprinting method, to have a preliminary idea of how bacterial diversity was distributed along the Victoria Valley. It amplifies the internal transcribed spacer (ITS) region between the 16S rRNA and the 23S rRNA genes, which is highly variable both in nucleotide content and length. Fluorescent primers are used and can be translated into peaks of abundance of the different fragment lengths in an electropherogram.

ITS regions from the 86 sampling sites were amplified according to the method adopted by Magalhães et al (2012). The OTU definition and statistical analysis were performed using Peak Scanner™ version 1.0 (Applied Biosystems) and the Primer 6 software package (version 6.1.11) (Clarke & Gorley, 2006) respectively. Fragments that differed by less or equal to 2 bp were considered identical, and fragments with FU (fluorescence units) below 50 were considered "background noise". Fragments less than 200 bp were removed, being too short ITS for bacteria. Then, values corresponding to peak areas were imported into the Primer 6 software package (version 6.1.11) (Clarke & Gorley, 2006). To study bacterial community structure, data were normalized using the presence/absence pre-treatment function and a lower triangular resemblance matrix was created using Euclidean distances and then examined using hierarchical cluster analysis. Dendograms were generated using the group average method and a Simprof test was applied to test differences between generated clusters.

Bacterial richness (S) and a diversity index (H') were calculated from the ARISA profiles to better address the ecological description of the bacterial community within samples. For these calculations, it was assumed that the number of peaks represented the species number (phylotype/genotype richness), and that the peak height represented the relative abundance of each bacterial species. The bacterial richness was expressed as the total number of unique OTUs (peaks) identified in each electropherogram. The Shannon-Wiener diversity index (H'), which takes into account both bacterial richness and abundance was also calculated using the PRIMER software (Clarke & Gorley 2006).

6. NGS analysis

6.1 DNA quality control

The six samples comprising the transect with increasing distance from a water source were selected for 16S rRNA amplicon sequencing. DNA quality was checked in a 1.5% electrophoresis gel stained with SYBR® Safe DNA Gel Stain (Invitrogen™). Additionally, bacterial 16S rRNA gene was amplified using the general Bacteria 27F/1492R primer set (Weisburg et al. 1991). PCR reaction consisted of 0.1 µM of each primer, 3.75 mM $MgSO_4$, buffer 10x, 0.1 mM dNTP's, and 2.5 U of Taqmed (5U/µl) for a final volume of 25 µl. The PCR program consisted of an initial denaturation step of 5 min at 94ºC, followed by 25 cycles of 45s at 95 ºC, 45s at 56 ºC and 90s at 72 ºC, followed by a final extension step of 10 min at 72 ºC. All reactions were performed in the Applied Biosystems Verifi 96 well thermocycler and the resulting ~1500 bp fragment was visualized in a 1.5% agarose gel (Grispo), in BioRad Molecular Imager Geldoc XR.

6.2. Pyrosequencing methodology

The 16S rRNA gene was amplified for the V3/V4 hypervariable region with barcoded fusion primers containing the Roche-454 A and B Titanium sequencing adapters, an eight-base barcode sequence, the forward primer 5'– ACTCCTACGGGAGGCAG-3' and the reverse primer 5'– TACNVRRGTHTCTAATYC -3' (Wang & Qian 2009). Two replicate PCR reactions were amplified from the same sample, in 20 µl reactions with Advantage Taq (Clontech) using 0.2 µM of each primer, 0.2 mM dNTPs, 5 U of polymerase, 6% DMSO and 2-3 µl of template DNA. The PCR conditions were 94 °C for 3 min, followed by 25 cycles of 94 °C for 30 s, 44°C for 45 s and 68 °C for 60 s and a final elongation step at 68 °C for 10 min. The amplicons were quantified by fluorimetry with PicoGreen (Invitrogen, CA, USA), pooled at equimolar concentrations and sequenced in the A direction with GS 454 FLX Titanium chemistry, according to manufacturer's instructions (Roche, 454 Life Sciences, Brandford, CT, USA) at Biocant (Cantanhede, Portugal).

6.3 Bioinformatic analysis

Raw pyrosequencing reads (FASTA files) were processed using the QIIME 454 pipeline (Caporaso et al. 2010b). QIIME (Quantitative Insights Into Microbial Ecology) is an open source package for comparison and analysis of microbial communities, primarily based on high-throughput sequencing data such as SSU rRNA generated from a variety of NGS platforms. The workflow consisted in several steps, and a general workflow is available for consultation in the Supplementary Information (Appendix 1; Fig.1).

In the demultiplexing step, the *split_libraries_output.py* was applied and consisted of an initial quality control, in which DNA sequence reads were filtered and assigned to their own samples (sampling sites T1-T6) by nucleotide barcode. Raw FASTA files, and a mapping file containing the specific barcode, served as input. Reverse primer was truncated and sequences with less than 200 nucleotides were discarded, as well as those with a quality score below 25.

In order to account for the known sequencing errors produced by the pyrosequencing technology, 454 datasets normally undergo an extra quality filtering process, which in our case was performed using Denoiser 0.9 (Reeder & Knight 2010).

Finally, in the *pick_de_novo_otus.py* workflow implemented on QIIME, UCLUST (Edgar et al. 2010) was used to cluster sequences and define OTUs at 97% sequence similarity. A representative sequence for each OTU was selected and aligned against the GreenGenes v12_10 (GG) database (DeSantis et al. 2006) using the PyNAST algorithm (Caporaso, et al. 2010a) and UCLUST for pairwise alignment (Edgar 2010). Sequences that satisfied the minimum length requirement and sequence identity (75% of median of the input FASTA file and 75% identity, respectively) were included in the alignment. The alignment of OTUs was filtered to remove gaps and hypervariable regions using the GG LaneMask (DeSantis et al. 2006). Then, the corresponding taxonomy was assigned to each OTU representative using UCLUST. The minimum percentage similarity to consider a database match a hit was 90% and the number of database hits to consider when making an alignment was defined to 3. A phylogenetic tree was constructed from the filtered alignment using the approximately maximum likelihood algorithm implemented in FastTree (Price et al. 2010), which was necessary for further beta diversity inferences.

6.3.1. Alpha-diversity metrics

Rarefaction curves were generated at 10 different sequencing depths, starting at a minimum depth of 10 sequences and ending at a sequencing depth corresponding to the maximum number of sequences for each sampling point. For each sampling depth, 1000 iterations were performed, to provide more robust results. In addition rarefaction curves for the chao1 richness estimator, Faith's PD (phylogenetic diversity) and the Shannon diversity index were also computed for each sampling point, using the QIIME *alpha_rarefaction.py* workflow.

6.3.2 Beta-diversity metrics

Beta diversity represents the comparison of microbial communities based on their composition. The main output of these comparisons are square distance matrices that can be visualized using Principal Coordinate Analysis (PcoA) and hierarchical clustering (UPGMA) methods. Differences between sites were calculated using the UniFrac metric (Lozupone & Knight 2005).

The *jackknifed_beta_diversity.py* workflow was used in order to generate weighted and unweighted UniFrac square distance matrices from the previously obtained OTU tables, depicting differences between communities. Then, a UPGMA tree was constructed for each distance matrix. To measure the robustness of this result in relation to the sequencing effort, a jackknifing analysis with 1000 iterations was performed, wherein 5769 sequences (75% of the total sequence number corresponding to the sampling point with lowest sequences) were chosen randomly from each sample. Distance matrices were generated for each of the 1000 subsamples of 5769 sequences for each sampling point and the resulting UPGMA trees were compared with the initial UPGMA tree representing the entire available data set. PcoAs of both weighted and unweighted UniFrac metrics reflecting final results were produced and visualized using Emperor (Vázquez-Baeza et al. 2013).

7. Alternative pipelines

There are several workflows available to analyze NGS data, and due to intrinsic characteristics of the bioinformatics tools and algorithms, it is not uncommon to observe some differences in overall results. Hence, and even though the main approach in this project was done using QIIME and all conclusions are drawn from the QIIME workflow, information

retrieved from two other pipelines used will also be mentioned. These pipelines used the same raw reads (FASTA files) as QIIME, but used different quality, filtering, trimming, assigning and clustering methods. The goal of using two other approaches is to establish comparisons between different pipelines, namely in terms of taxonomy attribution.

7.1. SILVA NGS

Raw sequence reads were processed by the NGS analysis pipeline of the SILVA rRNA gene database project (SILVAngs 1.2) (Quast et al. 2013). Each read was aligned using the SILVA Incremental Aligner (SINA SINA v1.2.10 for ARB SVN (revision 21008)) (Pruesse et al. 2012) against the SILVA SSU rRNA SEED and quality controlled (Quast et al. 2013). Reads shorter than 50 aligned nucleotides and reads with more than 2% of ambiguities, or 2% of homopolymers, respectively, were excluded from further processing. Putative contaminations and artefacts, reads with a low alignment quality (50% alignment identity, 40% alignment score reported by SINA), were identified and excluded from downstream analysis. After these initial steps of quality control, identical reads were identified (dereplication), the unique reads were clustered (OTUs), on a per sample basis, and the reference read of each OTU was classified. Dereplication and clustering was done using cd-hit-est (version 3.1.2; http://www.bioinformatics.org/cd-hit) (Li & Godzik 2006) running in accurate mode, ignoring overhangs, and applying identity criteria of 1.00 and 0.97, respectively. The classification was performed by a local nucleotide BLAST search against the non-redundant version of the SILVA SSU Ref dataset (release 119; http://www.arb-silva.de) using blastn (version 2.2.28+; http://blast.ncbi.nlm.nih.gov/Blast.cgi) with standard settings (Camacho et al. 2009). The classification of each OTU reference read was mapped onto all reads that were assigned to the respective OTU. This yields quantitative information (number of individual reads per taxonomic path), within the limitations of PCR and sequencing technique biases, as well as, multiple rRNA operons. Reads without any BLAST hits or reads with weak BLAST hits, where the function “(% sequence identity + % alignment coverage)/2" did not exceed the value of 93, remain unclassified. These reads were assigned to the meta group “No Relative" in the SILVAngs fingerprint and in Krona charts (Ondov et al. 2011).

7.2. Metabiodiverse

Raw pyrosequencing reads (FASTA files) were processed using Metabiodiverse, an automatic pipeline implemented at BIOCANT (Cantanhede, Portugal). In a first step, sequencing reads were assigned to the appropriate samples based on the respective barcode.

Reads were then, quality filtered to minimize the effects of random sequencing errors by elimination of sequence reads with <100 bp and that contained more than two undetermined nucleotides. Sequences in which the reverse primer was reached were additionally cut. Finally, sequences with more than 50% of low complexity regions, determined by DustMasker (Sogin et al. 2006) and chimera sequences, identified by UChime (Edgar et al. 2011), were discarded. The sequences were grouped by USearch (Edgar 2010) according to a phylogenetic distance of 3%, creating OTUs. Richness of population (rarefaction curves) and the diversity indices (Chao1) were calculated using the Mothur package (Schloss et al. 2009).

The taxonomy of each OTU was identified through a BLAST search against the Ribosomal Database Project II (RDP) database (Cole et al. 2009). The best hits were selected and subjected to further quality control. All sequences with an alignment coverage of less than 40% as well as those with an E-value greater than 1e-50 were rejected. Additionally, a bootstrap test was applied to the OTUs to identify the least common taxonomy level. Only the sequences with a bootstrap greater than 70% after 100 replicates, as obtained by seqBoot from Phylip package (Felsenstein 1989), were kept. The taxonomic assignment of the OTUs was completed with the attribution of the NCBI taxonomy identification number, which allowed the complete taxonomy construction of all identified organisms. Finally, for each taxon identified in the sample, the total number of sequences was summed up, providing the abundance of all identified organisms, for population statistics analysis.

IV. **Results**

1. Biological and geochemical attributes of Victoria Valley

From a total of 86 sampling stations over an area of 300 km^2, 9 were excluded due to technical difficulties in amplifying the ITS region. Therefore a total of 77 samples were used to get a preliminary idea of the distribution and ranges of both biological and geochemical variables in Victoria Valley. As biological variables, both species richness (S) and the Shannon index of diversity (H') were computed based on ARISA results, and rarefaction curves were plotted against the number of sampling sites (N). In addition, among 77 sampling sites, S ranges between 4 and 61, with a mean value of 31, and H' ranges from 0.972 to 4, with a mean value of 3 (Table 1).

The number of cells per gram of soil, derived from DAPI stained cell counts presents a range of four orders of magnitude (from 10^3 to 10^7), a tendency that is corroborated by measures of ATP activity, which ranges from 68.5 to 75189 RLU. In agreement to biological data, environmental parameters also show great variability, with pH ranging from 4.41 to 9.00, with an average of 7. Conductivity ranges from inexistence to 7050 µS/cm, with a mean value of 50 µS/cm (Table 1).

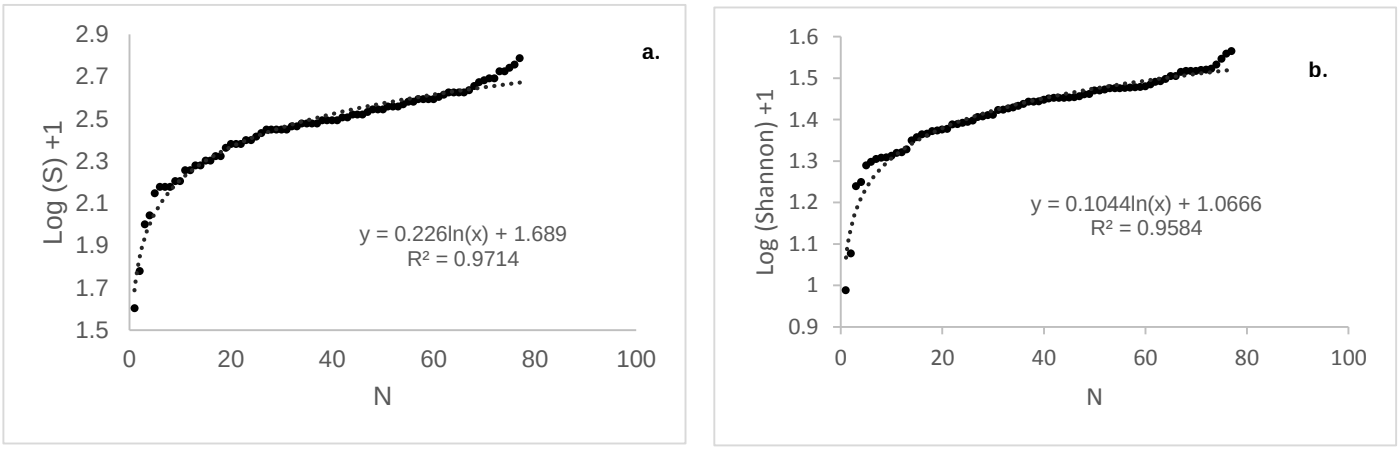


Figure 7. Graphical representations of log transformed species richness S (a) and species diversity H' (b). The value of each index was plotted against the number sampling sites (N) .

Correlation matrices using Pearson's R showed a positive correlation between S and H' which was expected, since the Shannon index takes species richness into account (Table 2). Also expected, was the positive and significant correlation between the ATP activity and the number of cells per gram of soil (Table 2) since generally, the higher the number of cells, the more ATP is generated.

In addition, ATP activity is also positively correlated with pH: higher ATP activity and consequently more cells were registered when pH gets more alkaline (Table 2). Interestingly the opposite correlation was verified for conductivity. Water activity is the variable that is stronger and positively correlated with H', ATP activity and number of cells (Table 2).

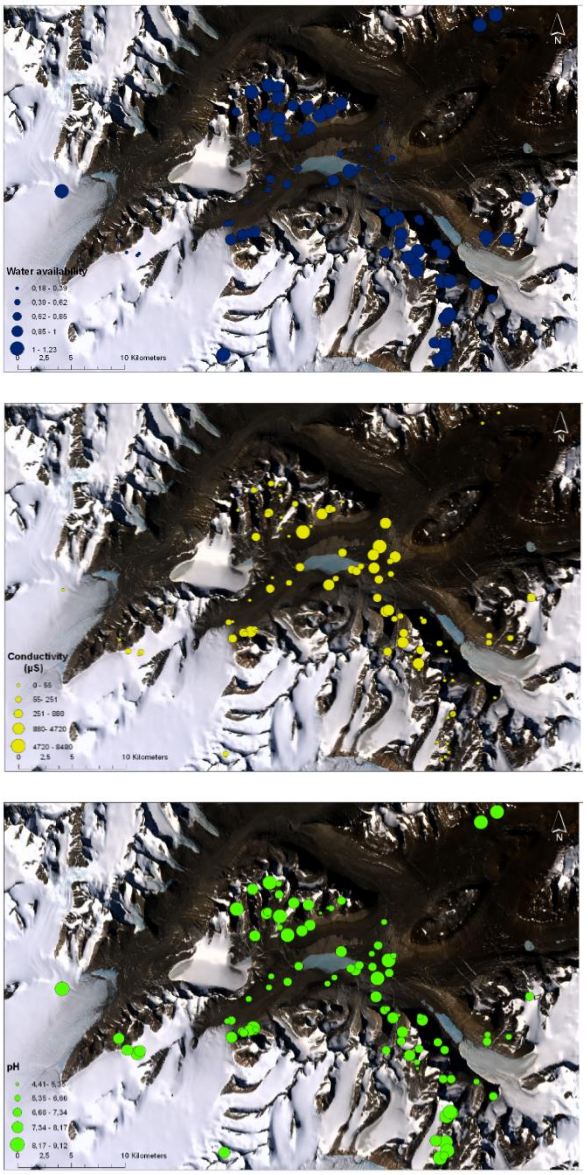


Figure 8. Spatial representation of water activity, conductivity and pH across the 86 sampling sites of Victoria Valley.

Table 1: Descriptive statistics for both biological and environmental variables collected in Victoria Valley (N=77). S = species richness; H' = Shannon index; A_w = water activity; ATP = ATP activity (RLU); Conductivity = µS/cm.

	S	H'	Cells/g	Aw	ATP	pH	Conductivity
Mean	31	3	2.58E+06	1	17323	7	50
Median	31	3	1.40E+06	1	12780	7	81
Mode	28	N.A	1.32E+05	1.05	N.A	7.6	0
Minimum	4	0.972	8.89E+03	0.18	68.5	4.41	0
Maximum	61	3.67	2.10E+07	1.23	75189	9.12	7010
SD	12	1	3.57E+06	0	16221	1	1162
Variance	1.42E-02	2.66E-01	1.27E-01	7.87E-13	2.63E-02	9.30E-01	1.35E-06

Table 2: Correlation matrix. Variables were normalized and the Pearson's R index of correlation was used to depict correlations between measured variables. In red are the R values that were considered to be significant ($p<0.05$). S = species richness; H' = Shannon index; A_w = water activity; ATP = ATP activity (RLU); Conductivity = µS/cm.

	S	H'	Cells/g	Aw	ATP	pH	Conductivity
S	-						
H'	0.8011	-					
Cells/g	0.0123	0.04	-				
Aw	0.1911	0.378	0.3615	-			
ATP	-0.1025	-0.029	0.2903	0.2483	-		
pH	-0.1416	-0.0846	-0.1522	-0.0848	0.2622	-	
Conductivity	-0.0426	-0.2318	-0.1096	-0.1944	-0.2685	-0.1626	-

2. Transect

2.1 Geochemical characterization

In addition to the variables previously described, in the six points belonging to the sampled transect, organic carbon content was also measured. All the biological variables regarding richness and diversity were derived from 16S rRNA gene pyrosequencing, which provides much more accuracy than ARISA fingerprint data. Due to the low number of sampling points (6), no correlation matrix was produced, since results would likely not reflect true tendencies. Regarding variable gradients across the transect, it is possible to observe a clear

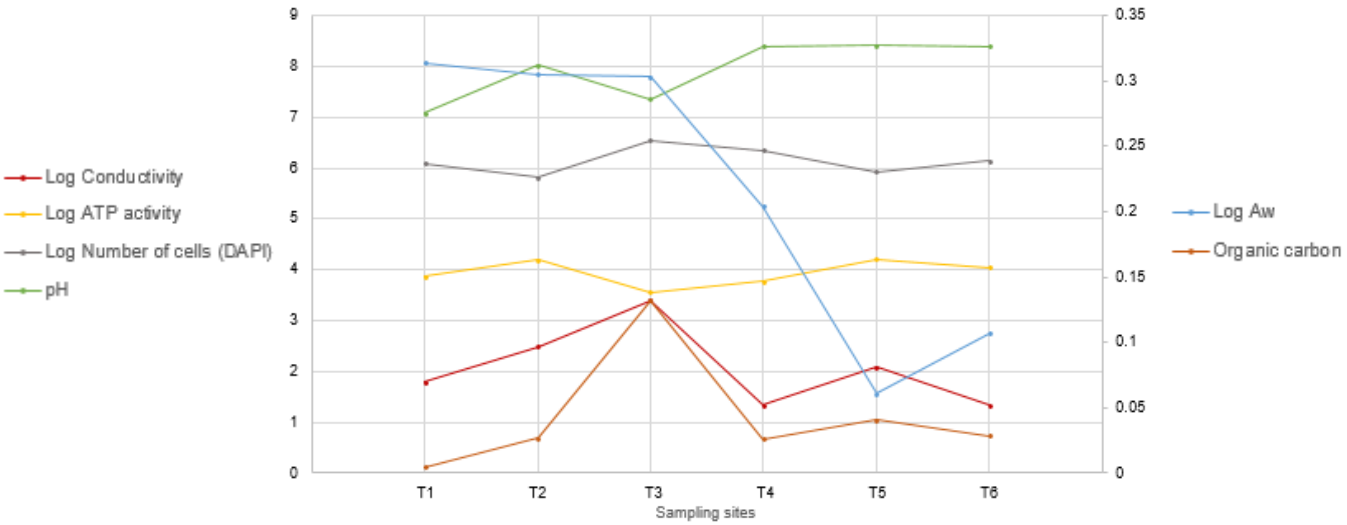


Figure 9. Variation of soil characteristics across the transect. Conductivity, ATP activity, number of cells and A_w were log transformed (log(x+1)), while percentage of organic carbon and pH were not. Conductivity, ATP activity, number of cells and pH are plotted against the primary axis y-axis (left), whereas A_w and percentage of organic carbon are plotted against the secondary axis y-axis (right). This was done in order to be possible to compare peaks of maximum and minimum values between the variables, as well as overall tendencies.

decrease in water availability, from the wetter sampling site (T1; A_w= 1.06), to the drier sampling site (T5; A_w= 0.15), which interestingly does not coincide with the site that is further away from the water source (Fig. 9). Regarding pH, it ranges from neutral (7.09) to moderately alkaline (8.41), and a slight increase as distance to water source increases (Fig. 9). Conductivity ranges from 21 to 2530 μS/cm and the maximum value corresponds to the sampling point where the percentage of organic carbon is also the highest (0.13%). In this same sampling point (T3) the number of cells per gram of soil is highest (3.46E+06), but ATP activity reaches minimum values (3638 RLU) (Fig. 9).

2.2 Bioinformatic analysis

Barcoded pyrosequencing analysis yielded 66947 raw sequences, which after the initial quality filtering/denoising/chimeric detection, decreased in number to 62789 (QIIME), 62319 (Metabiodiverse) and 64946 (SILVAngs) (Table 3). These sequences were clustered into 3897, 3724 and 5137 OTUs respectively with QIIME, Metabiodiverse and SILVAnsg (Table 4; Fig. 10). Even though the number of OTUs assigned by each pipeline differs, the overall pattern is similar, being T1 and T2 the richest ones, followed by T4 and T6. Then, Metabiodiverse and SILVAngs pipelines points T5 as richer than T3, whereas QIIME points to the opposite. Nevertheless, the number of OTUs is similar between T3 and T5.

Table 4. Number of available sequences after initial quality filtering, per sample, per pipeline.

	T1	T2	T3	T4	T5	T6	Total
QIIME	9273	7692	11893	11712	10357	11862	62789
Metabiodiverse	9422	7867	11181	11641	10366	11842	62319
SILVAngs	9865	8166	11423	12162	10834	12496	64946

Table 3. Number of unique OTUs per sampling site, per pipeline, after clustering

	T1	T2	T3	T4	T5	T6
QIIME	893	820	439	697	427	621
Metabiodiverse	944	803	357	671	398	551
SILVAngs	1209	1021	534	968	575	830

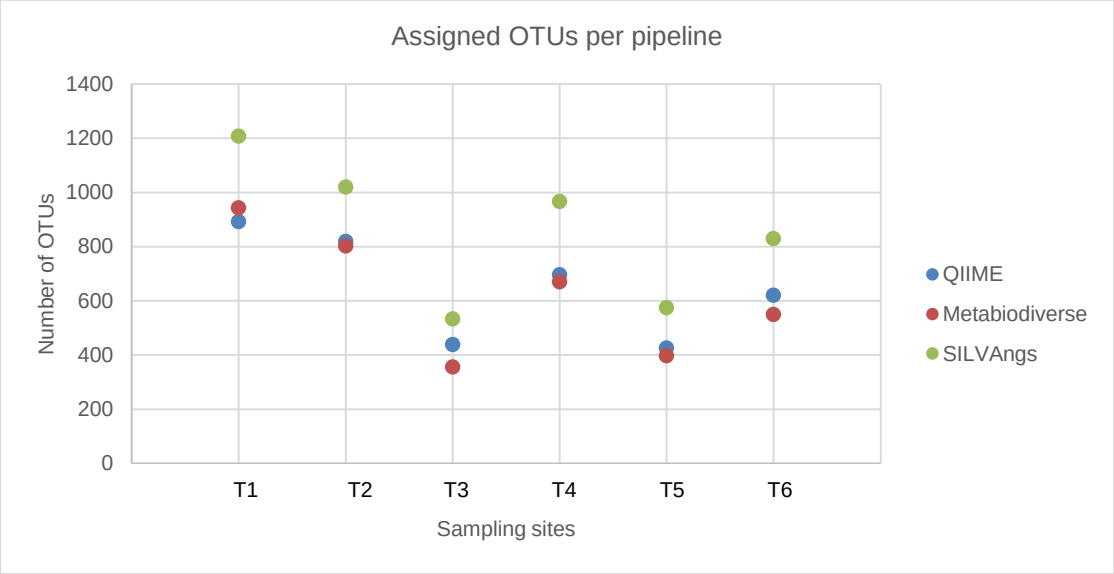


Figure 10. Graphical comparison of OTU attributions per sampling site, for each pipeline. Consistently, T3 is least diverse sample.

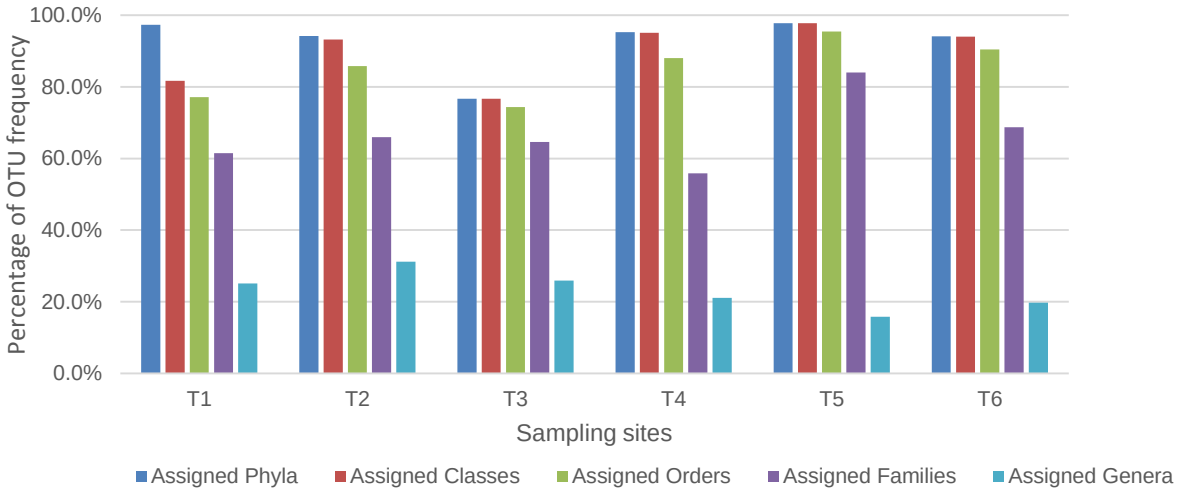


Figure 11. Percentage of OTU assignments per taxonomic group using QIIME: T3 is the sample with the highest percentage of unassigned OTUs, and there is a decreasing percentage of assignments, the deeper the taxonomix level.

A graphical representation of OTU taxonomic assignments per sample using QIIME is presented in Figure 11. Unassigned OTUs at the highest taxonomic level (phylum) ranged between 2.1% and 5.7%. However, in sample T3, the percentage of initially unassigned OTUs was much higher (23.3%). Also, the deeper the taxonomic level, the lower the percentage of assignments, due to marker accuracy.

2.2.1 Alpha diversity analysis

Rarefaction curves were computed for the observed species richness (number of unique OTUs per number of sequences), for the estimated species richness (given by the chao1 richness estimator), for the species diversity (given by the Shannon index of diversity), and for the phylogenetic diversity (given by Faith's PD). Comparing the curves of the true species richness (Fig. 12a) and the curves showing the estimation of species richness given by the chao1 index (Fig 12.b), it is possible to infer that: 1) T1 and T2 are the richest samples, and 2) T3 and T5 are the poorest samples, in terms of OTUs. In Figure 12a is possible to observe that sequencing depth did not fully capture the whole OTU richness in the samples, since particularly T1 and T2 are far from reaching a plateau, which would mean that no new

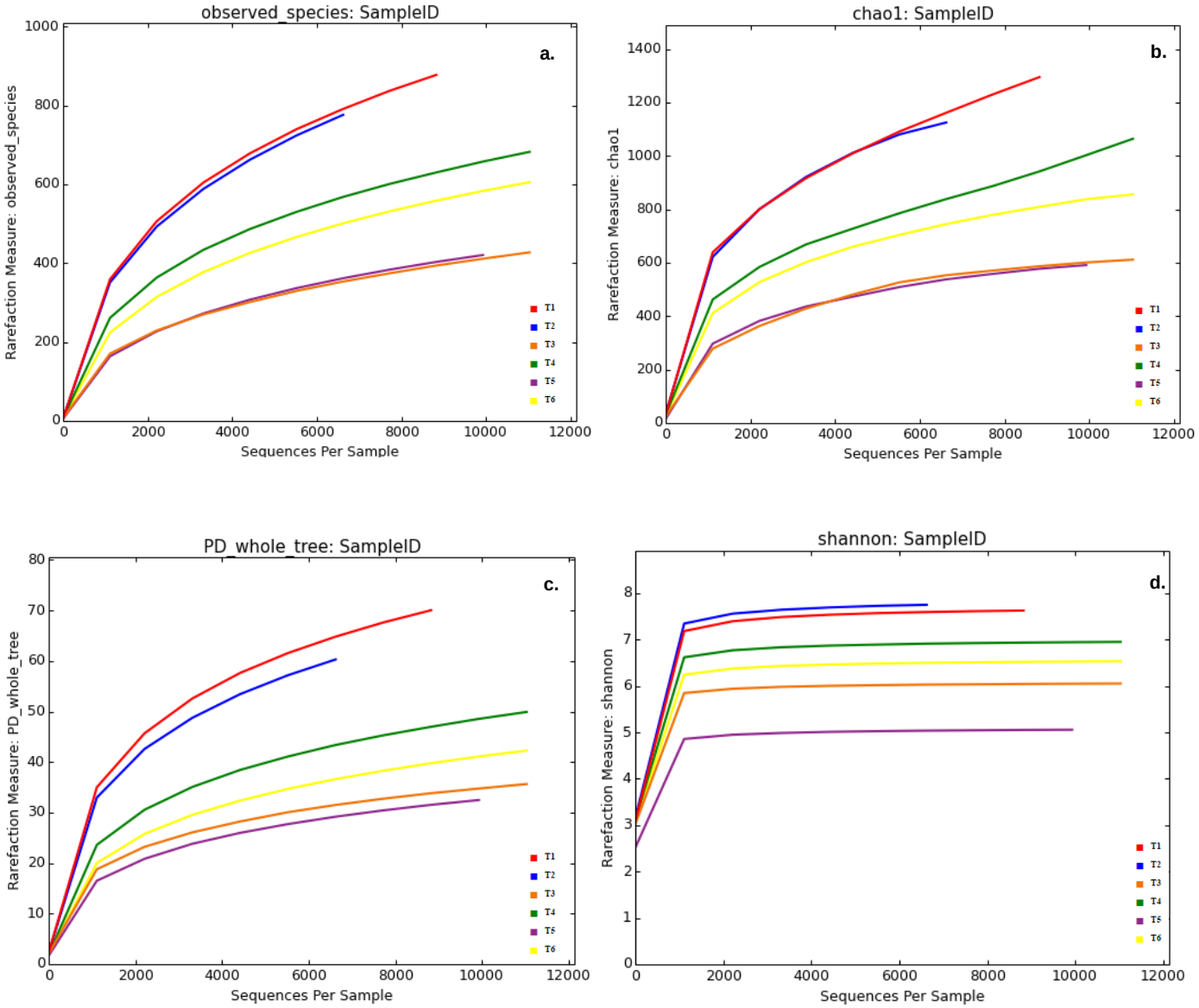


Figure 12. Rarefaction curves for different alpha diversity metrics for each sampling site. Observed species/OTUs (a); chao1 richness estimator (b); Phylogenetic diversity (Faith's PD) (c); Species diversity given by the Shannon index (d). General tendencies point to T1 and T2 as the richest and more diverse samples and T3 and T5 as the poorest. Additionally, sequencing depth did not successfully cover the full species richness in every sample (due to the fact that T1 and T2 do not reach a plateau in graphs a and b.

species/OTU would be found, even with an unlimited amount of sequences. The rarefaction curve for the chao1 richness estimator (Figure 12b) supports results from Figure 12a. Faith's PD represents the only alpha diversity index with a phylogenetic component and the observed pattern is similar to the rarefaction curves for the chao1 richness estimator and the observed diversity (Fig. 12c). According to the results obtained, T1 and T2 are pointed as the most diverse samples and are likely to possess species that are more phylogenetically different from each othe than T3 and T5, which are the less diverse, being probably composed of more similar taxa. Similarly to the rarefaction curves depicting the observed diversity and the chao1 richness estimator, the rarefaction curves of T1 and T2 regarding Faith's PD clearly do not reach a plateau (Fig. 12c).

Considering the rarefaction curves for species diversity given by the Shannon index, T1 and T2 are still the most diverse samples (with the highest values of the Shannon index), and T3 and T5 are the least diverse samples. The fact that in every sample, the values of the Shannon index clearly reach a plateau after >1500 sequences shows that the sequencing depth was successful in capturing the existing microbial diversity in our locations (Fig. 12d).

2.2.2 Beta-diversity analysis

The unweighted (Fig. 13) and weighted (data not shown) PcoA plots show similar clustering patterns, with the only difference being that the PC1 in the weighted Unifrac PcoA plot explained a higher percentage of variation than the PC1 in the unweighted Unifrac PcoA. T3 forms an independent cluster; then there is a cluster comprising T1 and T2, and a third cluster comprising T4, T5 and T6. Hierarchical clustering using both metrics showed that clusters were grouped at a very low UniFrac distance (<0.05), indicating that the soil samples presented a similar structure even though they were clustered apart.

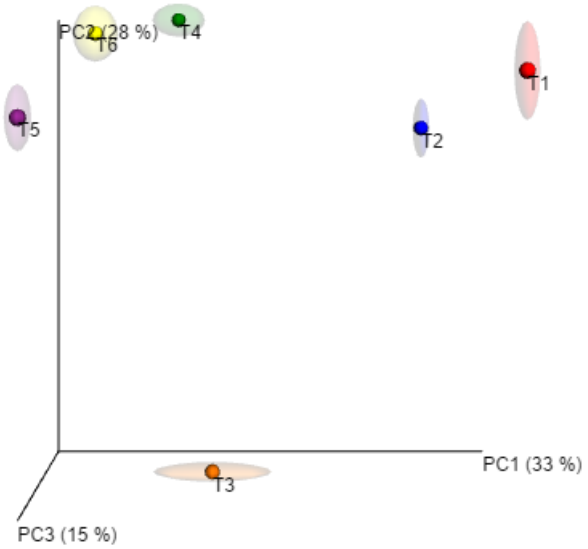


Figure 13. **Jackknifed PCoA plot using the unweighted UniFrac distance metric**. This metric accounts for presence/absence of taxa. The clusters were generated using a subset of 5769 sequences (75% of the number of sequences that constitute the least abundant sample) from each sampling site, for 1000 iterations. The positions of the points are the average for the jackknife replicates and ellipses were drawn around the mean values to represent the interquartile range.

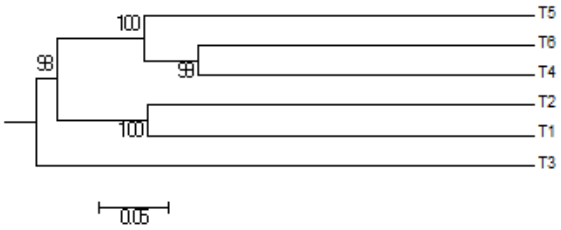
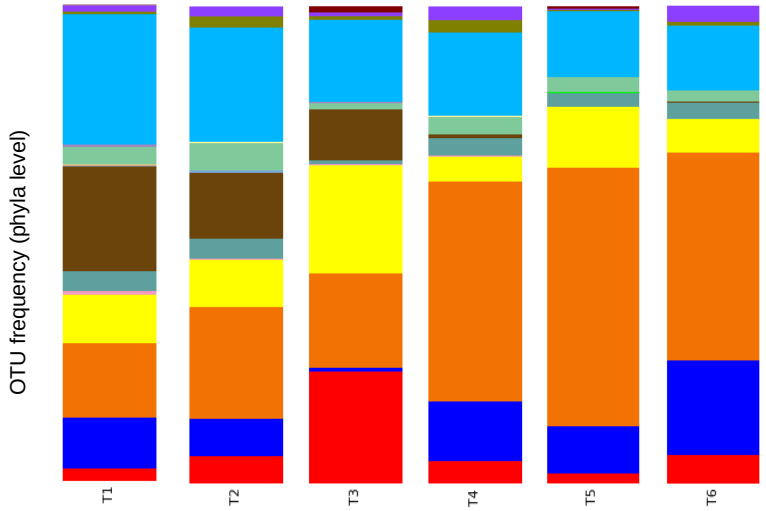


Figure 14. Unweighted UniFrac UPGMA tree. Numbers at branch points indicate the percentage of 1000 bootstrap re-samplings.

2.4 Detailed Bacterial taxonomic analysis

The relative abundance of each OTU at the phyla levels is depicted in Figure 15 where it is possible to observe the fractions containing unassigned OTUs in the bottom of the bar plots, in red. While there are phyla that are present in every sampling point at similar percentages of occurrence, other clearly change their frequency across the transect. In all the dataset, there are some dominant phyla, namely Actinobacteria (33.7%), Acidobacteria (10.2%), Bacteroidetes (11.3%), Cyanobacteria (7.9%) and Proteobacteria (18.9%). Also, the presence of members of phylum Deinococcus-Thermi was verified in all sampling points. Cyanobacteria are present in T1 (22.1%) and T2 (13.6%) and then their frequency falls to percentage levels of < 0.8% in the remaining sampling points, which are further away from the water source. There is also a notorious increase in the frequency of OTUs affiliated with the Actinobacteria phylum (from an initial prevalence of 15.48% in T1 to a prevalence of 54.13% in T5), with dominance of classes Acidimicrobia, Actinobacteria and Thermoleophilia (Supplementary Information: Appendix 2). This is particularly obvious with the Rubrobacteria class, for instances, which is represented by only one genus (*Rubrobacter*) which increases its frequency from (<0.1% to 2.5%) as the soils get drier (Supplementary Information: Appendix 2). Additionally, not only T5 has one of the highest OTU affiliation with Actinobacteria (54.13%), but 35.39% of the assigned OTUs belong to the Sporichthyaceae family (Supplementary Information: Appendix 4). The opposite pattern happens with Proteobacteria, which have an initial prevalence of 27.35% in T1 which decreases to 13.61% in T6). Alphaproteobacteria is the most abundant class within the Proteobacteria and the one that keeps a steady frequency along the sampling point, whereas Beta, Gamma and Deltaproteobaceria classes tend to diminish in frequency (Supplementary Information: Appendix 2).

Bacteroidetes are also present across the whole transect, with 6 representative orders, from which Saprospirales is the one with a more even distribution (OTU prevalence ranging from 1.1% to 5.5% (Supplementary Information: Appendix 3). However, in T3, not only there seems to exist a peak in the abundance in the overall frequency of Bacteroidetes (22.74% of all OTUs), but it is also possible to identify the Flavobacteriaceae and Cyclobacteriaceae families as major contributors, representing respectively 11.64% and 5.05% of the affiliated OTUs (Supplementary Information: Appendix 4).



	Phyla	Total	T1	T2	T3	T4	T5	T6
	Unassigned;Other	7.40%	2.71%	5.77%	23.34%	4.76%	2.19%	5.91%
	k__Bacteria;p__Acidobacteria	10.20%	10.63%	7.74%	0.96%	12.32%	9.87%	19.82%
	k__Bacteria;p__Actinobacteria	33.70%	15.48%	23.39%	19.72%	46.13%	54.13%	43.47%
	k__Bacteria;p__Armatimonadetes	0.00%	0.12%	0.08%	0.00%	0.05%	0.01%	0.04%
	k__Bacteria;p__BRC1	0.00%	0.01%	0.00%	0.03%	0.08%	0.07%	0.01%
	k__Bacteria;p__Bacteroidetes	11.30%	10.18%	9.89%	22.74%	5.09%	12.71%	6.96%
	k__Bacteria;p__Chlamydiae	0.00%	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%
	k__Bacteria;p__Chlorobi	0.20%	0.66%	0.34%	0.03%	0.20%	0.00%	0.02%
	k__Bacteria;p__Chloroflexi	3.20%	4.14%	4.21%	0.80%	3.71%	2.68%	3.58%
	k__Bacteria;p__Cyanobacteria	7.90%	22.05%	13.60%	10.72%	0.83%	0.24%	0.19%
	k__Bacteria;p__Elusimicrobia	0.10%	0.20%	0.09%	0.03%	0.04%	0.00%	0.02%
	k__Bacteria;p__FBP	0.00%	0.01%	0.05%	0.00%	0.00%	0.06%	0.02%
	k__Bacteria;p__Fibrobacteres	0.00%	0.02%	0.01%	0.01%	0.00%	0.00%	0.00%
	k__Bacteria;p__Firmicutes	0.10%	0.04%	0.38%	0.03%	0.00%	0.00%	0.01%
	k__Bacteria;p__GN02	0.00%	0.05%	0.00%	0.00%	0.00%	0.00%	0.00%
	k__Bacteria;p__Gemmatimonadetes	3.30%	3.70%	5.77%	1.39%	3.73%	3.25%	2.21%
	k__Bacteria;p__Nitrospirae	0.10%	0.60%	0.13%	0.02%	0.01%	0.00%	0.00%
	k__Bacteria;p__OD1	0.00%	0.00%	0.03%	0.03%	0.03%	0.00%	0.00%
	k__Bacteria;p__OP11	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
	k__Bacteria;p__Planctomycetes	0.00%	0.05%	0.03%	0.00%	0.02%	0.01%	0.04%
	k__Bacteria;p__Proteobacteria	18.90%	27.35%	24.14%	17.17%	17.62%	13.72%	13.61%
	k__Bacteria;p__TM6	0.00%	0.00%	0.01%	0.00%	0.01%	0.00%	0.00%
	k__Bacteria;p__TM7	1.20%	0.30%	2.17%	1.02%	2.52%	0.39%	0.65%
	k__Bacteria;p__Verrucomicrobia	1.70%	1.51%	2.05%	0.58%	2.77%	0.13%	3.32%
	k__Bacteria;p__WPS-2	0.00%	0.05%	0.00%	0.00%	0.01%	0.00%	0.01%
	k__Bacteria;p__WS3	0.00%	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%
	k__Bacteria;p__[Thermi]	0.40%	0.05%	0.10%	1.38%	0.09%	0.55%	0.12%

Figure 15. Frequency of phyla-affiliated OTUs per sampling point

2.5 Alternative pipelines

The same rarefaction approaches were performed in alpha diversity analysis from both Metabiodiverse and SILVAngs and results were similar to those obtained with QIIME, namely the fact that T1 and T2 are the richest and more diverse samples, whereas T3 and T5 were the poorest. Also, the rarefaction curves generated to assess diversity indexes within each sample, did not reach a plateau, especially in samples T1 and T2, suggesting the need of increasing the sequencing effort to fully capture the existing diversity.

The same pattern of phyla replacement (Actinobacteria-Proteobacteria) is observed with increasing distance from the water source, in the taxonomy reports of Metabiodiverse and SILVAngs. Also, T3 is considered the sampling point in which there is the highest percentage of unknown assignments, and the one where there seems to exist the highest percentage of OTUs affiliated with phylum Bacteroidetes. However, as it was predictable from the number of clustered sequences and total OTU counts per sampling point, there were some taxonomic assignments that were done in SILVAngs that were not seen both in the Metabiodiverse and in QIIME pipelines. Whereas QIIME identified 26 phyla (from which 9 were candidate divisions), and SILVAngs identified 29 (from which 12 were candidate divisions), Metabiodiverse only identified 15 phyla, all with culturable representatives.

V. Discussion

Environmental characterization *vs* microbial communities in Victoria Valley

Ice-free areas in continental Antarctica are scarce and patchily distributed, and the strong katabatic winds that characterize these regions are thought to be the major mean of dispersal of the few existing life forms (Marshall & Chalmers 1997; Nylen 2004).

The McMurdo Dry Valleys in the Transantarctic Mountains have been described has one of the few environments where abiotic factors (e.g., moisture, pH, conductivity) clearly drive the diversity and structure of the microbial communities (Wood et al. 2008; Pointing et al. 2009; Lee et al. 2012; Magalhães et al. 2012). This aspect together with the trophic simplicity of these systems makes Dry Valleys' soils perfect ecosystems to investigate the direct physical and chemical constraints on microbial biodiversity. The high range of microbial diversity observed among all our sampling sites (H' values ranging from 0.97 to 4) was matched by large differences in soil geochemical parameters. This is in agreement with the pronounced spatial heterogeneity of microbial communities that has been demonstrated to occur in other Dry Valleys at multiple spatial scales (Wood et al. 2008, Niederberger et al. 2008, Lee et al. 2012, Magalhães et al. 2012). In agreement, our ARISA data revealed large differences in bacterial diversity and community structure between samples that were even separated by short distances, indicating the occurrence of microbial population shifts even at small spatial scales. Indeed, the verified high environmental gradients that characterize Victoria Valley, and the resulting correlations between environmental variables and biological variables, namely species richness (S) and diversity (H') derived from ARISA results go accordingly to what has been found in other Dry Valleys (Cary et al. 2010), and strengthen the idea that environmental parameters contribute significantly to the distribution of soil microbiota.

Differences in soil geochemical properties, like ion composition, carbon content, water content, pH and nutrient availability are generally suggested as the major driving forces for the observed patterns of microbial distribution in the Transantarctic Mountains (Barrett et al. 2006b, Niederberger et al. 2008, Wood et al. 2008, Lee et al. 2012, Magalhães et al. 2012). Particularly soil moisture, together with soil conductivity are usually suggested as the prime (or the most important) determinants for species distribution in these extreme soil

environments (e.g. Porazinska et al. 2002; Barrett et al. 2004, 2006b, Niederberger et al. 2008; Cary et al. 2010). In agreement, our data indicated that the Victory Valley soils with lower conductivity and higher A_w favour higher bacterial diversity, prokaryotic abundance and corresponding metabolic activity (ATP).

The ranges of water activity and the corresponding richness and diversity indexes that characterize each location are particularly interesting to point out. According to literature, an A_w of 0.6 is generally considered to be the lowest limit for the existence of microbial life (Grant 2004; Bolhuis et al. 2006;). However, our measurements of A_w ranged from 0.15 to 1.23 and we still managed to detect the presence of bacteria (by both ARISA and pyrosequencing), suggesting that there are organisms in these extreme environments that thrive even below the current lower A_w threshold, indicating that organisms that inhabit the polar desert soils region are tolerant or adapted to very low moisture conditions. Similarly to other valleys, namely Beacon and Battleship Promontory, Victoria Valley does not have a direct marine influence, since it is not a coastal valley. It has a medium height of approximately 1000 m, being part of the inland-mixed microclimatic zone, which is one of the three microclimate zones that characterize the Dry Valleys (Marchant & Head 2007). Mean temperatures tend to decrease with height, therefore Victoria Valley is also characterized by lower mean temperatures than other coastal valleys. Hence, when compared with other studied valleys (Miers, Taylor, Upper Wright), Victoria Valley seems to possess harsher environmental constraints to the existing microbiota.

Water availability *vs* bacterial diversity

Spatial heterogeneity in physical, chemical and biological properties of soils allows the proliferation of diverse microbial communities. In this context, the Dry Valleys' soils fit perfectly, as they present a mosaic of glacier, ice-covered lakes, ephemeral streams and arid soils, being the closest of Earth's analogs to Mars. Bacterial community fingerprinting results indicated that there are communities that thrive, in Victoria Valley, and that are adapted to the strong environmental constraints which are characteristic of this valley and of the Dry Valleys in general. In order to evaluate the influence of one particular factor – water availability – in the bacterial community dynamics, we sampled a transect with a gradient of water availability and assessed how the bacterial communities changed (in terms of taxa richness and diversity inferred by pyrosequencing the 16S rRNA gene) as distance to water increased.

The best characterized microniche communities in the Dry Valleys are the cryptoendolithic microbial consortia, which are mostly found in the fine-grained sandstones and in coarse-grained quartzites and limestones. The combination of porosity (which provides interstitial spaces) and translucence (which facilitates photosynthetic activity) allows microbial growth below and between the rock surface. Microorganisms are thus protected from physical abrasation by wind-blown sand and at the same time are buffered from thermal fluctuations. Hypolithic communities may also be found in Dry Valleys' soils. These communities thrive around the undersides and around the margins of translucid rocks such as quartz and marble. The translucence of the rock allows the transmission of light, enabling the survival of phototrophs and facilitating the development of more complex communities. Hypolithic communities may also represent local "hotspots" for microbial diversity and productivity (Cary et al. 2010).

The heterogeneity of the existing taxa across the transect may be indicative of a switch in habitat from a cryptoendolithic-like environment, comprising T1 and T2 which were the most diverse locations, and the ones closer to the water source, to open arid soils, which comprise locations T4, T5 and T6 and correspond to the locations which are further away from the water source, but are not necessarily the least diverse samples. The location between these two environments – T3 – could represent a hypolithic-like environment, since it is the location with the highest percentage of organic carbon. The reason for this assumption that there could be a change in environments is based in beta-diversity inferences and subsequent clustering. Both weighted and unweighted UniFrac PcoA plots showed the same clustering pattern, indicating the existence of three clusters: one comprising T1 and T2; one comprising T4, T5 and T6; and another one comprising T3. The Unweighted UniFrac is a qualitative metric that

is sensitive to factors that affect the presence/absence of bacterial communities such as salinity, water and/or pH, whereas the weighted UniFrac is a quantitative metric that is sensitive to factors that affect relative abundance of taxa, such as seasonal changes, nutrient availability, response to pollution, etc (Lozupone & Knight 2005; Lozupone et al. 2007). Both metrics yield different, complementary results and applying both metrics can provide insights into the nature of community differences (Lozupone et al. 2007). In our case, PC1 of the weighted UniFrac metric seemed to better explain the observed clustering pattern, suggesting that taxa abundance could be a factor differentiating the microbial communities across the transect.

T3 is the last location with values of $A_w > 1$ across the transect, and is the sample that is least diverse, according to all pipelines. However, it is also the one with the highest percentage of unknown assignments, which could boost the diversity indexes. In addition, T3 represents the location with the highest percentage of organic carbon, conductivity, and number of cells, and also the one with the lowest values of ATP activity. Therefore, microorganisms that were not identified in our analysis could be present in this location and, in association with members of phylum Bacteroidetes (which also present their highest frequency in this location), be responsible for the lowest values of ATP, and maximum number of cells and percentage of organic carbon that characterize this location. In fact, heterotrophic members of phylum Bacteroidetes are known to be involved in the degradation of complex polymers (Thomas et al. 2011). Interestingly, QIIME detected the existence of 16 bacterial phyla in this location, tagged 22% of the sequences as “unassigned” and pointed phylum Bacteroidetes as dominant, comprising 22.7 % of the assigned sequences. The SILVAngs pipeline identified 18 bacterial phyla in this location (an increase justified by the inclusion of more Candidate phyla), tagged 9% of the sequences as “non-relative”, and also pointed phylum Bacteroidetes as dominant, comprising 32% of the assigned sequences. The Metabiodiverse pipeline identified 13 bacterial phyla in this location and showed the same tendency as the SILVAngs pipeline regarding the dominance of Bacteroidetes in this location (31.5% of the assigned sequences). In addition, it was the only pipeline that identified members of the domain Eukarya, particularly members of phyla Bacillariophyta in this location, providing stronger evidence that this location could harbor hypolithic communities, which would be responsible for higher values of local productivity. A way to confirm this hypothesis would be to analyze the sediments in terms of composition, and assess the existence of, for example, translucid rocky sediments.

Cryptoendolithic environments are known to harbor an extensive and varied microbial community, most of which dominated by Cyanobacteria, Proteobacteria (particularly Alphaproteobacteria), and *Deinococcus spp*. This goes accordingly to the taxa found in

locations T1 and T2, which are the ones closer to the water pond and the locations with the highest values of A_w. Alpha diversity metrics for these two locations showed that, in these locations, the existing OTU richness was still far from completely captured with only this sequencing effort, as the rarefaction curves for all metrics (except the rarefaction curves for the Shannon index) were far from reaching a plateau. Cyanobacteria represented respectively 22.1% and 13.6 % in T1 and T2, whereas Alphaproteobacteria represented 16 and 13.8%. Members of the highly radiation and dissecation resistant phylum Deinococcus-Thermi were ubiquitous across all locations, showing a higher prevalence in drier and open soils, in which radiation is stronger and water content is lower.

As the soils get drier, the environment changes to an open soil, in which Actinobacteria are known to proliferate and dominate over other bacterial taxa (Ventura et al. 2007). This Gram-positive phylum is ubiquitous in soils and comprises members which are dissecation-resistent, such as members of genus *Rubrobacter*, whose presence was detected in higher frequencies, in locations T4, T5 and T6. In addition, our results showed that three orders of Actinobacteria dominated drier locations, with a clear dominance of Actinomycetales over Acidomicrobiales and Solirubrobacterales. Soil actinomycetes are important degraders of organic matter (Hayakawa & Nonomura 1987), and are important in maintaining environmental stability. They have also attracted considerable scientific interest due to their ability to produce a range of bioactive metabolites, including antibiotics, with genus *Streptomyce*s being the largest antibiotic-producing genus. Previous studies have reported actinomycetes in several Antarctic ecosystems. Strains affiliated with genera *Corynebacterium*, *Micrococcus*, *Nocardia* and *Streptomyces* have been reported in the McMurdo Dry Valleys (Cameron et al. 1972), as well as a considerable number of novel psychrotolerant actinobacterial species belonging to the family Microccocaceae, from other Antarctic locations (Liu et al. 2000). Members of the known psychrotolerant *Arthrobacter* genus have also been reported in Antarctic ecosystems (Gupta et al. 2004).

Our results confirm the prevalence and dominance of actinomycetes in the hyper-arid soils of Victoria Valley, as strains affiliated with genera *Arthrobacter*, *Aeromicrobium*, *Friedmanniella*, *Nocardioides* and *Pseudonocardia* were detected. Interestingly, the highest percentage of occurrence of actinomycetes was present in the driest location (T5), was not classified down to the genus level, and all we were able to infer was that sequences were affiliated with family Sporichthyaceae. Taxonomically, this family is sometimes included in the sub-order Frankiales, and according to literature, is only composed of one spore forming genus – *Sporichthya* – that encompasses two described species, one of which – *S. polymorpha* – reported at very low frequencies in Antarctic soils (Babalola et al. 2009).

Therefore, the fact that sequences affiliated with this family represent such a big fraction (35.4 %) of the sequences in T5, may be indicative of a still undescribed species. The SILVAngs pipeline also tagged 34% of the sequences in T5 as an uncultured strain within sub-order Frankiales, and Metabiodiverse tagged 30% of the sequences in T5 as unassigned members of order Actinomycetales. Uncultured and undescribed species of actinomycetes that thrive in such extreme conditions represent an untapped source of genetic diversity and ultimately a possible source of novel bioactive metabolites.

NGS bioinformatic analysis

Pyrosequencing of 16S rRNA genes allows for an in-depth characterization of complex microbial communities. Hypervariable regions of the 16S rRNA gene are flanked by conserved sequences, which enable the design of “universal” PCR primers that can amplify a designated region from a large number of different bacterial species (Nikolaki & Tsiamis 2013). The accumulation of known polymorphisms in the conserved regions means that the coverage rates of some primers are declining (Jonasson et al. 2002), which can cause problems in using widely accepted primers, since they fail to recover a high percentage of bacterial species in uncultured environmental samples. A recent study from Wang & Qian (2009) has predicted all the potential primers for the bacterial and archaeal 16S rRNA gene and concluded that their position is largely consistent with those of known primers, but the average coverage rate is higher than known primers, alerting to the fact that there is a need to update the existing primers sets as new taxa are added to the databases, from which primers are designed on the first place (Wang & Qian 2009). Our primer set amplified the V3/V4 hypervariable region of the eubacterial 16S rRNA gene. Even though it seems to provide a good coverage (Wang & Qian 2009), we cannot exclude the possibility that it failed to recover some taxa (namely those which are not abundant), neither can we exclude the possibility of having slightly different results if we had used another primer set.

Current sequencing technologies provide many reads per run and researchers found out that it is cheaper, less time consuming and equally valid to combine multiple samples in a single run. This process is named “multiplexing” and, in the case of Roche’s 454 pyrosequencing technologies, is achieved through the application of a pyrosequencing-tailored nucleotide barcode design (Parameswaran et al. 2007). By performing this step, DNA can be randomly sequenced at once and be later assigned according to the respective sample, in the data processing step.

Large datasets are generated after a high throughput sequencing run, and there is the need to have adequate computational tools that can address this kind of sequencing data. Another issue that influences final results is the methodologies which are used in data processing, namely the OTU picking method/clustering algorithm, the alignment algorithm, the taxonomic classification algorithm, and probably the most important: the reference database to which taxonomic assignments are made against (McDonald et al. 2012) (Table 4). A robust universal reference taxonomy is a necessary aid to the interpretation of high-throughput sequence data from microbial communities and even though taxonomy based on the 16S rRNA gene is the most comprehensive and widely used in microbiology today, it has yet to reach its full potential because numerous microorganisms belong to taxa that have not yet been characterized and also because numerous sequences that could be reliably classified, remain unannotated. This is the case of GenBank in which two thirds of the 16S rRNA gene sequences are only classified to the domain level and is probably the most widely consulted 16S-based taxonomy. In order to overcome these limitations, several dedicated 16S databases emerged, such as Greengenes (DeSantis et al. 2006; McDonald et al. 2012), the Ribosomal Database Project (Cole et al. 2009), and SILVA (Quast et al. 2013), which encompass a higher proportion of environmental sequences and have to be manually curated.

Table 5. Summary of the major differences between pipelines. References for each algorithm were previously mentioned.

	Alignment	Clustering	Taxonomy	Ref. Database
QIIME	PyNAST	UCLUST	RDP Classifier	Greengenes v12_10
Metabiodiverse	USEARCH	USEARCH	BLAST	RDP Database II
SILVAngs	SINA	CD-HIT	BLAST	SILVA SSUrRNA SEED

Ideally, all these databases would synchronize with each other, so that results would be the same, independently of the reference database used. However, this does not happen, and there is the constant need to add, update and manually curate new sequences in each database, and since this is not done automatically, some taxonomic discrepancies can be observed, since one particular database may be more complete than the other.

One of the goals of this study was to account for this discrepancy, and three reference databases were used (one for each pipeline), in order to see if there would be a taxonomic gap resulting from the use of one reference database, that could be filled by other/others. This was the case with the SILVAngs pipeline, that used the SILVA SSUrRNA SEED database as reference, which is one of the databases with the highest number of entries (Huse et al. 2008) and the one that assigned the presence of more candidate phyla, in contrast with the Metabiodiverse pipeline, that used the RDP Database II, and did not assign any sequence to

candidate phyla. The Greengenes reference database is implemented by default in QIIME and is pointed by some authors as preferential over other databases (Navas-Molina et al. 2013). Regarding taxonomic classification, there are also many algorithms available (Cole et al. 2009; Edgar 2010; Fu et al. 2012), and the most commonly employed practical approach is the naïve Bayesian method developed for the RDP (Cole et al. 2009). This algorithm has proven its utility and has sustained considerable popularity since its introduction, and is also the recommended taxonomic classifier by the QIIME team, with a confidence value of 0.8 (Navas-Molina et al. 2013). The clustering algorithms also differed between the three pipelines, and differences were mostly in computational speed. Whereas CD-HIT is more computationally demanding, the USEARCH and UCLUST algorithms are faster, use less memory and have improved sensitivity (Edgar 2010). The ideal clustering algorithm should balance the inclusion of sequences into an OTU that are within a specific genetic distance, while excluding those that are greater than that distance. The algorithm used by UCLUST, which is the average neighbor algorithm, seems to outperform the weighted, nearest and furthest neighbor algorithms (being the latter the one used by both CD-HIT and USEARCH), according to recent studies (Schloss & Westcott 2011). The alignment method also differed, and even though the recently released SINA alignment algorithm has been suggested as more accurate than PyNAST (Pruesse et al. 2012), the latter was the chosen one, given the limited algorithm choices implemented in QIIME (which were Infernal, ClustalW, Muscle and Mafft algorithms). Nevertheless, since its release (Caporaso et al. 2010a), PyNAST has been widely used in several studies (e.g. Chu et al. 2010; Kuczynski et al. 2012).

One important aspect of our analysis with QIIME, which did not take part in the analysis made by the other two pipelines (SILVA and Metabiodiverse), was that after the initial quality filtering and trimming, sequences were denoised. This is particularly important in the case of pyrosequencing data, because of the chemistry of the PCR reaction that produced characteristic sequencing errors which are mostly imprecise signals for longer homopolymer runs, which could further give rise to false OTUs. Nevertheless, this denoising step is still discussable, particularly in cases where the dataset is considered small (< 50000 sequences), since it can also lead to the exclusion of rare OTUs. As our raw dataset was composed of approximately 60000, we opted for denoising our data, but the analysis of a non-denoised dataset against our results, could be a way to assess whether or not we lost information by performing this step. Additionally, as opposed to the other pipelines, we did not check for the existence of chimeric sequences. We did, however, use the Greengenes reference database, which is chimera-checked, and therefore potential chimeras in our data would likely be tagged as “unassigned”.

In conclusion, the choice of the NGS data processing parameters is flexible and should be adapted to the type of biological question, as for example, some databases are known to be more complete, regarding some particular taxa, than others. The QIIME pipeline was chosen as the primary tool to process our data, since it is a flexible and easy pipeline to work with, which results from the implementation of the most up-to-date algorithms in a single piece of software, making it easier to understand what each algorithm does, and to manipulate particular parameters within the analysis.

VI. Conclusions

This study combined community fingerprinting and next generation sequencing methodologies to assess the direct effect of the high environmental gradients that characterize Victoria Valley, in terms of the existing bacterial communities. Similarly to what has been observed in other valleys, Victoria Valley is characterized by intrinsic prokaryotic communities that can thrive below thresholds that are considered to be life-limiting. The highly marked differences observed in the biological parameters (bacterial diversity, metabolic activity and total cell counts) measured in 77 sampling stations were found to be intrinsically related with the strong environmental gradients that characterized Victoria Valley. These findings demonstrated once more that microbial dispersion, diversity and abundance in the Antarctic Dry Valleys are dictated by strong physical and chemical pressures, being abiotic factors the main controls of microbial communities' development in these habitats. The effect of one particular environmental parameter – water availability – was evaluated and a clear shift between microbial communities was registered. This shift was characterized by a pattern of phyla replacement, as distance to the water source increased, likely resulting from a shift in habitat from cryptoendolithic-like to open arid soils. In addition, one particular intermediate location showed the highest rate of unassigned sequences, the highest frequency of occurrence of members of phylum Bacteroidetes, and the highest percentage of organic carbon which can reflect of the existence of more complex communities in this potentially hypolithic location.

Future work will consist in an increase in the sampling and sequencing effort, since the full taxa richness was not captured with this analysis. In addition, new transects in other valleys will also be sampled, to potentiate inter-valley comparisons and a more accurate knowledge on how water availability dictates microbial communities' development in these extreme cold deserts. Geological insights regarding the type of sediment/rock we are dealing with will also be of the utmost importance, because different sediments harbor distinct microbial communities. Also, the choice of a universal primer that amplifies simultaneously for both Bacteria and Archaea will also be a hypothesis, since recent studies are emerging, reporting the existence of the latter domain in Antarctic ecosystems.

Chapter 2. Diversity of ammonia-oxidizing microorganisms in the Transantarctic Mountains

I. Background

1. Nitrogen cycle

Nitrogen is the most abundant element on the Earth's atmosphere and an essential component of proteins and nucleic acids (Thamdrup 2012). Nevertheless, most of the nitrogen is under the form of atmospheric nitrogen (N_2), it cannot be directly used by living beings and therefore, has to be reduced/fixed by capable organisms in order for the element to be available in downstream biologic processes. The nitrogen cycle (Fig. 16) is composed of several pathways, in which microorganisms play a key role (Francis et al. 2007).

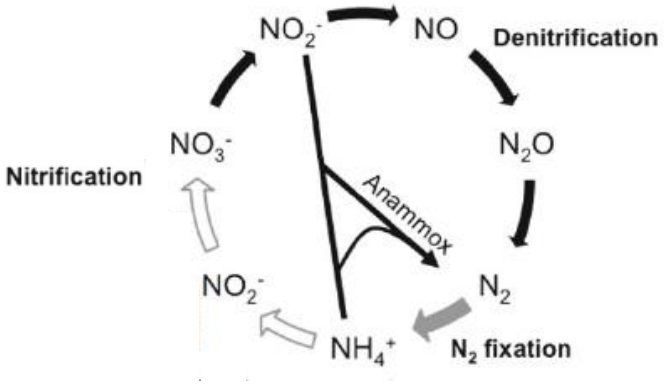


Figure 16. Illustration of the key processes that comprise the nitrogen biogeochemical cycle. In grey is the ammonia-oxidizing step of nitrification, followed by the oxidation of nitrite to nitrate. Adapted from Monteiro et al. (2014)

Nitrogen fixation consists on the fixation of atmospheric nitrogen (N_2) to produce ammonia (NH_4^+), which can then be used in downstream biological processes. This pathway specific enzymes – *nitrogenases* – which exist in Cyanobacteria and symbiotic and non-

symbiotic diazotrophs such as members of the *Rhizobium* genus and the Azotobactereaceae family respectively (Wagner, 2011).

Another step is the nitrification process, which can be sub-divided in two pathways: the aerobic oxidation of NH_4^+ to nitrite (NO_2^-), which is the main focus of this chapter, and the subsequent oxidation of NO_2^- to nitrate (NO_3^-). The oxidation of NH_4^+ is mediated by ammonia-oxidizing Bacteria (AOB) (Kowalchuk & Stephen 2001; Arp et al. 2007) and Archaea (AOA) (Schleper & Nicol 2010; Hatzenpichler 2012) and the oxidation of NO_2^- is mediated by members of the bacterial phylum Nitrospira, and genera *Nitrobacter, Nitrococcus and Nitrospina* (Francis et al. 2007; Hayatsu et al. 2008). Initially, ammonia-oxidation was thought to be an exclusively aerobic process, but evidence that it also occurred under anoxic conditions – Anammox - changed this assumption (Francis et al. 2007). In this alternative pathway, members of the Planctomycetes phylum fully convert NH_4^+ and NO_2^- to N_2 and H_2O, using NO_2^- as electron donor instead of molecular O_2 (Hayatsu et al. 2008).

The following step – denitrification – is carried out by representatives of the *Pseudomonas* and *Clostridium* genera and is an anaerobic reaction that consists on the reduction of NO_3^- to N_2 (Francis et al. 2007). This step is particularly important, since the presence of NO_3^- can have undesirable consequences such as soil detriment and algal blooms, respectively in agricultural and wastewater treatment systems. Along with the production of N_2, nitrous oxide N_2O is also produced as a result of denitrification. This gas is known to be a strong greenhouse gas, involved in climate change due to its interaction with the ozone layer (Jung et al. 2014; Kozlowski et al. 2014).

Anthropogenic activities such as the addition of nitrogen-based fertilizers and burning of fossil fuels are known to have a great impact on the nitrogen cycle, by increasing the available nitrogen in natural systems. This phenomenon has increased exponentially with industrialization and has severe consequences in terrestrial and marine ecosystems, leading to an unbalanced distribution of nutrients, changes in carbon storage, habitat degradation and changes in the food-web structure, and increased acidification in freshwater ecosystems (Jung et al. 2014).

2. Diversity of ammonia-oxidizing microorganisms

The first step of the nitrification pathway involves the autotrophic oxidation of NH_4 to NO_2^-, with hydroxylamine as an intermediate. This process requires two different enzymes: ammonia monooxygenase (AMO) and hydroxylamine oxidoreductase (HAO) (Arp et al. 2007) (Fig. 17). Ammonia oxidation reactions yield a low energy budget, and therefore, microorganisms that are responsible for this step are generally slow growers and difficult to

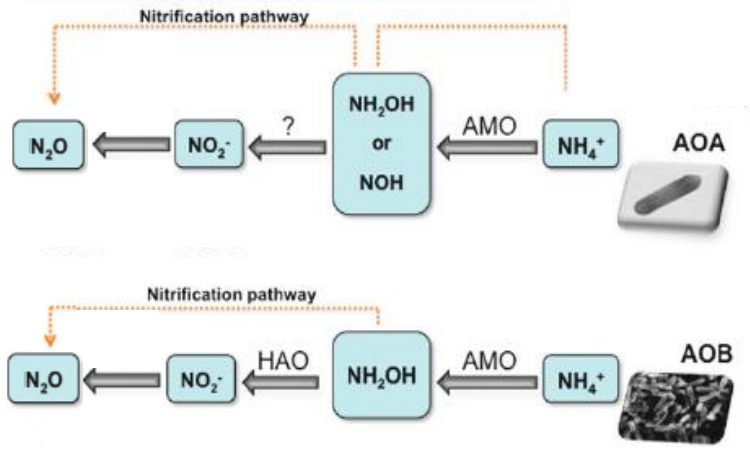


Figure 17. Key processes of nitrification by both Archaea (AOA) and Bacteria (AOB) ammonia-oxidizers (adapted from Monteiro et al., 2014).

grow in pure cultures (Kowalchuk & Stephen 1997). Unlike the other pathways of N cycle that are carried out by many different kinds of microbes, the ability to oxidize NH_4^+ is less broadly distributed among prokaryotes, and until recently, only Bacteria were thought to mediate this step. However, a series of recent discoveries derived from the development of culture independent methods have revolutionized this view. First, the role of anaerobic ammonia oxidation was confirmed (Mulder et al. 1995), and secondly, the analysis of metagenomic libraries from sea water (Venter et al. 2004) and soil (Treusch et al. 2005), revealed the existence of putative genes involved in ammonia oxidation in genomic fragments derived from uncultivated Crenarchaeota.

The cultivation of the first marine ammonia-oxidizing archaeon, *Nitrosopumilus maritimus* (Könneke et al. 2005) and the further isolation and cultivation of ammonia oxidizing Archaea from other mesophilic and thermophilic environments (Hatzenpichler et al. 2008; De La Torre et al. 2008; Tourna & Stieglmeier 2011) have resulted in a broader knowledge regarding the ecology and physiology of ammonia-oxidizing organisms, as well as the relative

role of bacterial and archaeal ammonia oxidation in a wide range of natural ecosystems (Francis et al. 2005; Leininger et al. 2006; Magalhães et al. 2009; Monteiro et al. 2014).

The 16S rRNA gene was the first phylogenetic marker to be used to assess the existing diversity of both AOA and AOB in natural ecosystems and has significantly contributed to our knowledge regarding their current phylogeny (Junier et al. 2010). However, its main pitfall is that it is not necessarily related to the physiology of the target organisms. In addition, in most of the cases, ammonia-oxidizing microorganisms are not numerically dominant and therefore, the 16S rRNA gene-based methods may not have sufficient sensitivity to distinguish all the existing representatives (Junier et al. 2010). Therefore, the use of the functional genes that are under selective pressure, such as the AMO alpha subunit encoding *amoA* gene*,* were suggested as complementary markers to assess the diversity of ammonia-oxidizing microorganisms. This particular gene was revealed to have a strong potential for fine-scale differentiation of closely related ammonia-oxidizers (Rotthauwe et al. 1995; Rotthauwe et al. 1997). The AMO enzyme is encoded by a three-gene operon (*amoC-amoA-amoB*) which exists in 2-3 copies in bacterial ammonia-oxidizers of the Betaproteobacteria class, and only in single copy in bacterial ammonia-oxidizers of the Gammaproteobacteria class (Arp et al. 2002; Arp et al. 2007). The AMO operon in Archaea varies from the traditional arrangement seen in Bacteria as well as in copy numbers, ranging from one to three. (Spang et al. 2012).

Taxonomically, bacterial ammonia-oxidizers fall into two lineages within the phylum Proteobacteria. The first is affiliated with the Betaproteobacteria class and includes genera *Nitrosospira*, *Nitrosomonas*, and one representative of genus *Nitrosococcus* (*N. mobilis*) which, according to different authors, can also be classified as *Nitrosomonas mobilis* (*e.g.* Purkhold et al. 2003; Koops et al. 2006) ;The second lineage is affiliated with the Gammaproteobacteria class and includes all the remaining members of genus *Nitrosococcus* (Purkhold et al. 2000; Koops & Pommerening 2001; Purkhold et al. 2003; Koops et al. 2006).

Regarding Archaea, the analysis of "Candidatus *Cenarchaeum symbiosum*" genome led to the proposal of a novel phylum of Archaea – Thaumarchaeota – since, until then, AOA were classified as mesophilic Crenarchaeota (Preston et al. 1996; Könneke et al. 2005). The further availability of new isolates, genomes and enrichment cultures reinforced the uniqueness of this phylum (Brochier-Armanet et al. 2008; Hatzenpichler et al. 2008; Spang et al. 2012; Stahl & de la Torre 2012; Pester et al. 2012). Within the Thaumarchaeota, there are four defined clusters: the marine group 1.1a (also known as *Nitrosopumilus* cluster) that includes genus *Nitrosopumilus* and *Cenarchaeum symbiosum*, the soil group 1.1b (also known as *Nitrososphaera* cluster), which includes the members of genus *Nitrososphaera*, and the *Nitrosotalea* and *Nitrosocaldus* clusters, with only one culturable representative each

(*Nitrosotalea devanaterra*, and *Nitrosocaldus yellowstonii*, respectively) (Stahl & de la Torre 2012).

3. Nitrogen biogeochemistry in the Transantarctic Mountains

Antarctic microbial communities are often unusual and intrinsically interesting because they have been subjected to long periods of isolation with relatively low levels of disturbance. In addition, there is a pressing need to understand Antarctic microbial ecosystems before they may be irreversibly damaged by human activities and climate change (Vincent 2004).

The McMurdo Dry Valleys constitute the largest relatively ice-free region of continental Antarctica and were formed by the advances and retreats of glaciers through the coastal ranges of the Transantarctic Mountains. Even though they were initially classified as “sterile” and “unable to hold life”, the Dry Valleys are now known to harbor surprisingly high biodiversity levels, most of which microbial (Pointing et al. 2009; Cary et al. 2010). However, the dynamics of the N biogeochemical cycle are still not fully understood. To our knowledge, there have been only two studies focusing on the existence of critical genes involved in nitrification in the Dry Valleys’ soils (Chan et al. 2013; Magalhães et al. 2014). The other few existing studies from these areas are limited to lakes (Voytek & Ward 1995; Voytek et al. 1999), and to the less extreme sub-Antarctic and maritime peninsula (Yergeau et al. 2007a; Jung et al. 2011). Recent studies focusing on the Archaea domain in the Dry Valleys, consistently showed a high prevalence of 16S rRNA gene sequences (80-99%) belonging to Crenarchaeota/Thaumarchaeota, affiliated with the soil group 1.1b (Ayton et al. 2010; Richter et al. 2014), representing increasing evidence of the importance of Archaea in maintaining nitrification rates on the Dry Valleys. Environmental variables are seen as major rulers of the distribution and relative abundance of general microbial communities (Lee et al. 2012; Richter et al. 2014), as well as AOA/AOB communities (Magalhães et al. 2014). Variables such as substrate concentration (i.e NH_4^+) may play an important role in the distribution and composition of nitrifying communities, and there have been reports of an AOA numerical dominance over AOB in low NH_4^+ and oligotrophic environments (Martens-Habbena et al. 2009; Zhalnina et al. 2012). Additionally, high C/N ratios, pH, and conductivity have been strongly correlated with higher levels of AOA and/or AOB abundance, in these extreme environments (Magalhães et al. 2014).

The terrestrial landscape in the Dry Valleys is dominated by oligotrophic mineral soils and extensive exposed rocky surfaces, where high physicochemical heterogeneity namely altitude, surface geology, salinity, water availability, pH, and temperature represent strong

constraints to the existence of life. Hence, the existing biodiversity is thought to form simple trophic relationships (Hogg et al. 2006; Barrett et al. 2006b; Cary et al. 2010). In this study, we present an extreme case about how environmental variables dictate the relative diversity and abundance of AOA and AOB under environmental conditions that we believe are near the limit of their distribution and of their ability to function.

II. Goals

In this study, we investigated the distribution and phylogenetic relationships of Bacteria and Archaea ammonia oxidizers in a range of Antarctic soil environments across the Transantarctic Mountains, based on DNA molecular approaches (DNA extraction, amplification, cloning and sequencing). A functional gene (*amoA*) coding for a subunit of an enzyme involved in the ammonia-oxidation process was targeted in both AOA and AOB and diversity was related with the extreme environmental variables that characterize each location. We believe that this study is critical to understand the dynamics of the nitrogen cycle in permanently cold environments, and to provide insights and a better understanding on how those microbial communities respond to changes in polar ecosystems.

III. Material and methods

1. Sample collection and site description

Sampling campaigns were integrated in the New Zealand Terrestrial Antarctic Biocomplexity Survey (NZTABS - http://nztabs.ictar.aq/) and included four Antarctic expeditions: the Darwin-Hartherton Glacier region of the Darwin Mountains (79° 55´S, 157° 35´E) in December 2007; Miers Valley (78°6′S 164°0′E) and Beacon Valley (77°83‟S, 160°66‟E,) in December 2006; Upper Wright Valley (77°10′S, 161°50′E) and Battleship Promontory (76°54'S 160°55'E) in January 2008, and finally, Victoria Valley (77° 23´S, 162° 00´E) in January 2013.

The Darwin-Hatherton Glacier region is characterized by xerous soils in which factors such as terrain age, glaciation history and soil geochemistry are pointed as the major drivers of microbial existence (Magalhães et al. 2012). Samples were collected near Lake Wellman, and at Junction Spur. Miers Valley is a coastal, low altitude valley with a high C/N ratio

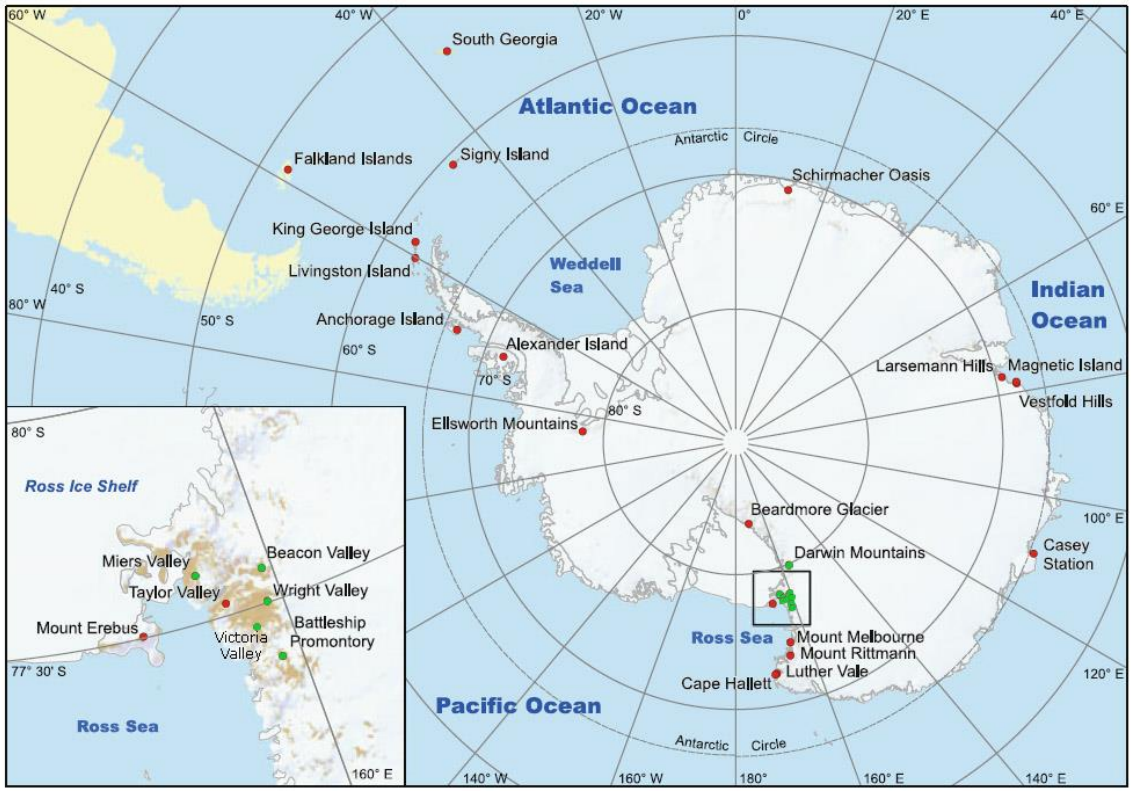


Figure 18. Map of Antarctica, emphasizing ice-free regions. Locations assessed in this study are marked with green dots. Adapted from (Bottos et al. 2014)

which has been the setting of many microbial studies involving bacterial communities (e.g. Wood et al. 2008; Lee et al. 2012). Beacon, Victoria and Upper Wright Valleys have a higher mean altitude and are characterized by strong katabatic winds, low C/N ratio, lower temperatures and high soil electrical conductivity, presenting a set of unfavorable conditions to the existence of life (Lee et al. 2012). Battleship Promontory is also a high altitude valley with temporary liquid water in snow melt points, leading to higher moisture contents, and lower electrical conductivity, therefore providing milder conditions for the existence of bacterial communities (Lee et al. 2012).

In Beacon, Upper Wright, Battleship Promontory and Miers Valleys, two 50m perpendicular transects were defined from which four sampling points comprising $1m^2$ (A-D) were defined at the edges of each transect. A scoop of soil was collected aseptically from the top 2 cm at the four corners of the defined 1 m^2 area, combined in a sterile Whirl-Pak (Nasco International Inc., Fort Atkinson, WI, USA) and stored at -80 ºC as soon as possible, for later analysis. In the 2007 expedition to the Darwin-Hartherton Glacier and also in Victoria Valley (in 2013), in each sampling point an area of $1m^2$ was defined and five soil samples for each sampling point were collected (from the edges and the center of the 1 m^2 area), combined a sterile Whirl-Pak (Nasco International Inc., Fort Atkinson, WI, USA) and stored at -80 ºC for later analysis.

All the procedures involved in the sampling campaigns, as well as all the *in situ* measurements were performed and managed by NZTABS in-field multidisciplinary teams. All necessary and appropriate precautions were made in order to avoid anthropogenic or cross-site contaminations. Aliquots of soil samples were shipped to Oporto's University and analyzed on behalf of this project.

2. DNA extraction

Samples were extracted using a modification of the CTAB (Cetyltrimethylammonium bromide) extraction protocol (Barrett et al. 2006b). One gram of homogenized soil from each sample site was weighted and phosphate (100 mM NaH_2PO_4) and SDS buffer were added (100mM NaCl, 500 mM Tris, pH 8.0, 10% SDS). After agitation, vortexing and recovering of the supernatant, the CTAB solution was added and samples were incubated for 30 minutes. Afterwards, they were washed twice with chloroform isoamyl alcohol (24:1), centrifuged and the supernatant was retrieved. Then, 7M of ammonium acetate buffer was added, samples were centrifuged, the supernatant was retrieved and 0.54 volumes of

isopropyl alcohol were added. After centrifugation, and recovery of the supernatant, the samples were submitted to an overnight incubation. In the following day, samples were washed in 70% ethanol, eluted in 25 µl of nuclease-free water and stored at -80 ºC. Two replicates were extracted for the same sampling site

3. PCR amplification of Archaeal and Bacterial *amoA* gene

To assess AOA and AOB *amoA* gene diversity, two soil samples of each sampled location were selected to build clone libraries. Archaeal *amoA* amplification was performed using PuReTaq Ready-to-Go PCR Beads (GE Healthcare), and the Arch-*amoA*F/Arch-*amoA*R primer set described by Francis et al. (2005). Bacterial *amoA* amplification was also performed using PuReTaq Ready-to-Go PCR Beads (GE Healthcare), as well as the *amoA*1F/*amoA*2R' primer set described by Rotthauwe et al. (1997) and Okano et al. (2004). Total reaction volume was 25 µl and 1 µl of template DNA was used. PCR was performed in an Applied Biosystems Verifi 96 well thermocycler under the following conditions: an initial denaturation step of 5 minutes at 94 °C, followed by 9 pre-cycles of denaturation at 94 °C for 30 seconds, annealing at 65 °C for 30 seconds and extension at 72 °C for 30s. Next, 35 cycles of 30 seconds at 94 °C, 30 seconds at 55 °C and 30 seconds at 72 °C were performed, and finally an extension step of 7 minutes at 72 °C was carried.

4. Cloning and RFLP profiling

PCR products were run in a 1.5% electrophoresis gel, stained with SYBR® Safe DNA Gel Stain (Invitrogen™). The bands were excised and purified using the GF-1 Gel DNA Recovery Kit (Vivantis) according to manufacturer's instructions. Amplicons were cloned using the TOPO TA Cloning Kit (Invitrogen™) according to manufacturer's instructions. The only change performed in the cloning protocol was the decrease of the vector DNA, salt solution, and competent cells to half the volume suggested in the kit. Ten to thirty colonies were randomly selected from the library of each gene, for each sampling site. Plasmids were isolated using the GeneElute plasmid miniprep kit (Sigma) and DNA concentrations of purified plasmids were determined fluorometrically using Qubit 2.0 (Invitrogen™). Insert size was verified by digesting 0.6–0.8 µg of DNA at 37 ºC for 2 h with 10 U of *EcoRI* (Sigma-Aldrich, Portugal) and a total of 399 clones of AOA and AOB were screened for similarity by RFLP analysis after digestion with MspI (Promega, Europe). After RFLP screening, a total of 109 clones were selected for sequencing at the STABVIDA Sequencing Facilities (Lisbon, Portugal).

5. Phylogenetic and statistical analysis

All retrieved sequences were quality checked and trimmed using Sequence Scanner (Applied Biosystems). Then, they were imported and aligned in MEGA 6.0 (Tamura et al. 2013) using the Clustal W algorithm (Thompson et al. 1994), and a p-distance matrix was generated and further imported to Mothur 1.10.2 (Schloss et al. 2009) where OTUs were defined respectively with a clustering cut-off of 95% and 85% for AOA, and 96% and 80% for AOB, using the average neighbor algorithm. A consensus sequence for each OTU was generated using Jalview 2.0 (Waterhouse et al. 2009) and they were blasted against the National Center for Biotechnology Information (NCBI) database to obtain the closest published sequences. The first different eleven hits (E=0) on NCBI for each consensus sequence were selected for Neighbour-Joining (NJ) tree construction, which was performed in MEGA 6.0 (Tamura et al. 2013) with 1000 replicates to produce bootstrap values.

IV. Results

1. Archaeal *amoA*

RFLP was used as a method for clone screening, in order to have a preliminary idea of the number of existing phylotypes, so that only the representatives of a given phylotype would be sequenced. The major pitfall of this approach is the fact that the profiles generated by enzyme digestions may or may not correspond to the real profiling, which is given by DNA sequencing. For AOA, a total of 177 clones were produced in the cloning step, corresponding to 5 distinct RFLP phylotypes (data not shown). Representatives of these phylotypes (n=36) were sequenced and clustered at 85% and 95% respectively. This yielded 3 and 5 OTUs respectively and after comparison with RFLP phylotypes, there was a correspondence between the attributed phylotypes and the nucleotidic OTUs clustered at 95%. The same did

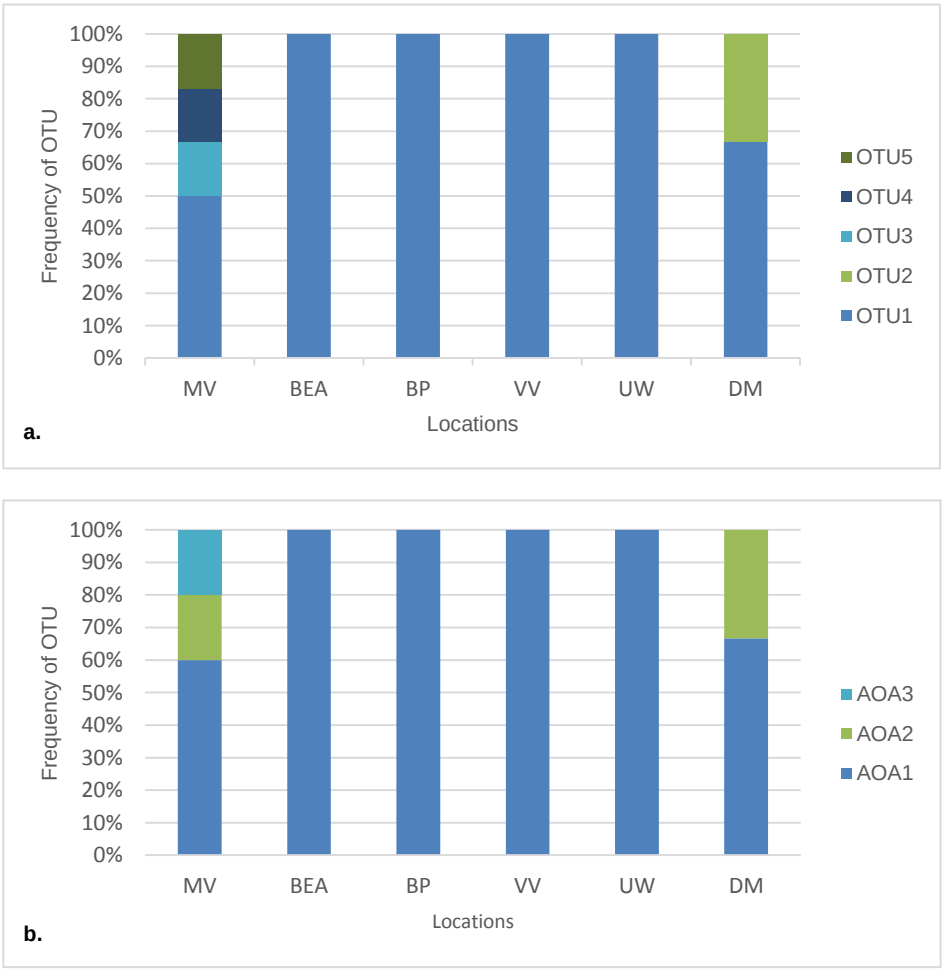


Figure 19. Number of retrieved archaeal OTUs with different clustering cut-offs, per location. a) a 95% clustering cut-off yielded 5 archaeal OTUs, and b) a 85% clustering cut-off yielded 3 archaeal OTUs. MV- Miers Valley; BEA – Beacon Valley; BP- Battleship Promontory; VV – Victoria Valley; UW - Upper Wright Valley; DM – Darwin Mountains

not happen with the clustering cut-off of 85% and therefore future inferences will be based only on the nucleotidic profiles. The overall AOA *amoA* gene diversity was low, with only five and two OTUs identified with a clustering cut-off of 95% and 85% respectively (Fig. 19). With the 95% cut-of, OTU1 was the most abundant, comprising 83% of the clones (n=30), whereas the remaining 17% of the sequences were distributed within OTU2 (n=3), OTU3 (n=1), OTU4 (n=1) and OTU5 (n=1). Miers Valley was the location with the highest number of OTUs, with four out of five, followed by the Darwin Mountains. In the remaining locations (Beacon, Battleship, Wright and Victoria Valley), only one *amoA* OTU was found. In terms of OTU prevalence, OTU1 was the most abundant, being present in all sampling sites. OTU2 was exclusive of the Darwin Mountains, whereas OTU3, OTU4 and OTU5 were exclusive of Miers Valley, making it the most diverse location, in terms of archaeal *amoA*. The consensus sequences of all OTUs showed respectively 97-98%, 88-92%, 92-93%, 97-99%, 88-93% nucleotide sequence similarity to sequences recovered from other environments such as arable soils (Glaser et al. 2010), wetlands (Wang et al. unpublished), pristine forest soils (Szukicks et al. unpublished), alkaline soils (Song et al. 2014) and Tibetan plateau soils (Wu, X. K, unpublished; Zeng, J., unpublished; Xie et al.unpublished), respectively. Nevertheless, when blasting the OTUs against archaeal *amoA* gene sequences of culturable representatives, nucleotide similarity percentages were significantly lower, ranging from 74-80%.

With a clustering cut-off of 85%, only 3 AOA OTUs were retrieved: AOA1 (which resulted from a merging of OTU1 and OTU3) was the most abundant, comprising 83% of the clones (n=30), whereas the remaining 17% of the clones were distributed within AOA2 (which resulted from a merging of OTU2 and OTU5) (n=4) and AOA3 (n=1), that corresponded to former OTU4. Again, Miers Valley stood out as the most diverse, possessing all the retrieved OTUs. The Darwin Mountains were considered the second most diverse location, with two out of the three retrieved OTUs, whereas all the other locations presented only one OTU (AOA1). The consensus sequences for all three OTUs showed respectively 97-99%, 87-97% and 97-99% sequence identity to sequences recovered from other environments such as tundra soils (Daebeler et al. 2012), arable soil (Glaser et al. 2010), pristine forest soils (Szukicks et al. unpublished) and Tibetan plateau soils (Wu X. K, unpublished; Zeng, J., unpublished; Xie et al.unpublished), respectively.

The neighbor-joining (NJ) phylogenetic trees depicting the relationships among archaeal *amoA* gene sequences shows that the general differences among these sequences are small in terms of substitutions per site (Fig. 20; Fig. 21). Taxonomically, AOA are known to form four clusters: *Nitrosopumilus* (marine 1.1a cluster); *Nitrososphaera* (soil 1.1b cluster),

Nitrosocaldus (formerly clustered in the marine 1.1a cluster), and *Nitrosotalea* (composed of only one member *N. devanaterra*). In both trees, it is possible to distinguish the presence of archaeal representatives of the 1.1a (marine) and 1.1b (soil) subgroups, as well as the *Nitrosocaldus* cluster (Fig. 20; Fig. 21). All the retrieved OTUs seem to be affiliated with the soil 1.1b cluster, composed of members of the genus *Nitrososphaera*, namely *N. gargensis*, *N. viennensis* and *N. evergladensis*. There also seems to exist two well supported sister groups in the AOA trees: one comprising OTU5 / OTU2 (AOA2) and the other one comprising OTU1 / OTU3 / OTU4 (AOA1 and AOA3). Within the latter group, OTU4 is the one that is more closely related to the *Nitrososphaera* cluster. OTU1 and OTU3 are closely related to soils of pristine environments, meadows soils and wetland sediments. OTU5 and OTU2 form a more basal group and are closely affiliated with clones from alkaline and glacier soils. Interestingly, even with such a decrease in clustering cut-off, OTU4 (AOA3) remained separated from all the other OTUs, with no close culturable affiliations.

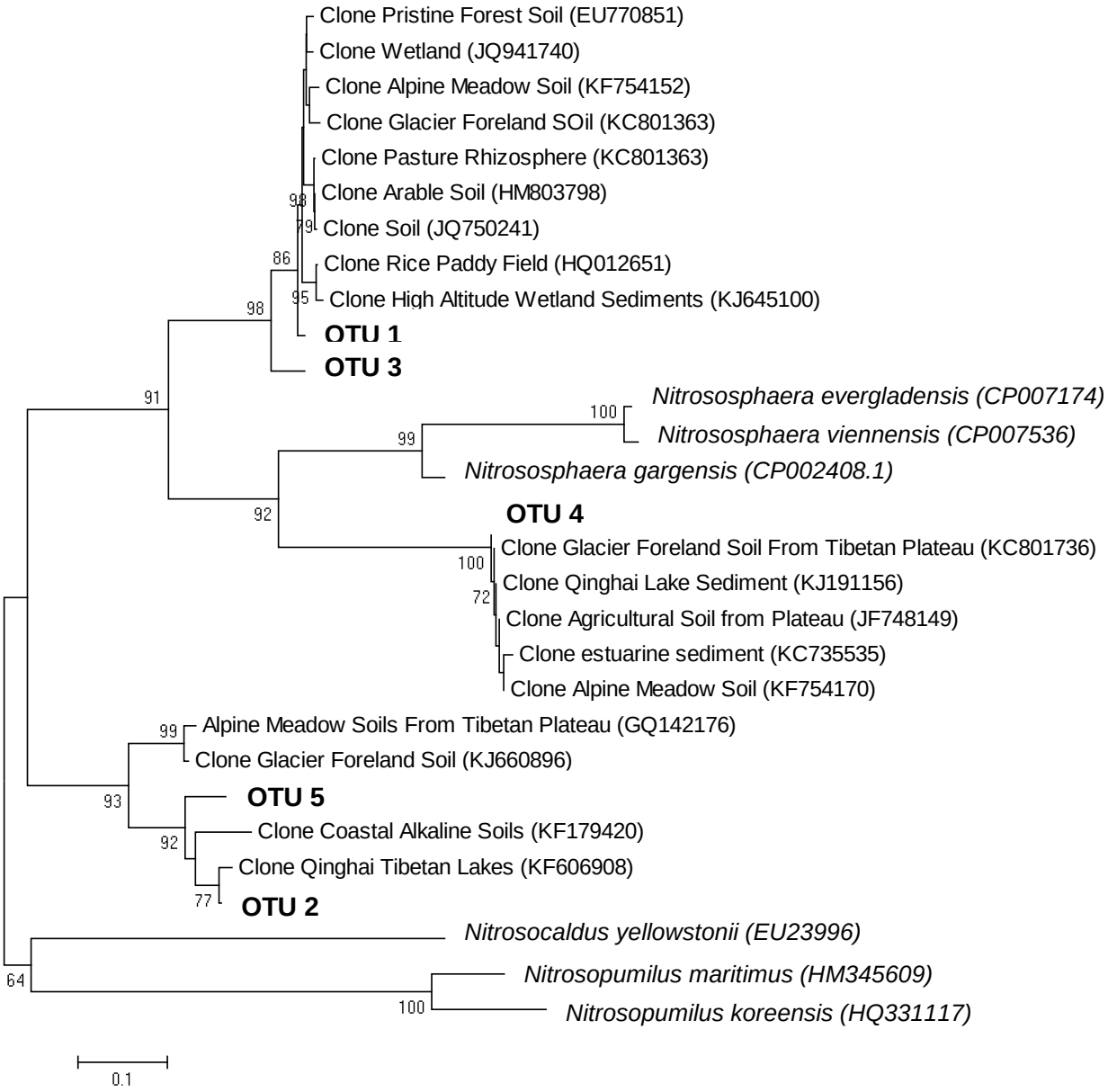


Figure 20. Unrooted *amoA*-based phylogenetic tree of ammonia-oxidizing Archaea with a 95% clustering cut-off. The 630-bp gene fragment using the archaeal primer set proposed by Francis et al. (2005) was used for phylogeny inference which was done in MEGA 6.0 (Tamura et al. 2013), using the NJ treeing method and the T92 + I + G nucleotide substitution model, with 1000 replicates to provide bootstrap support. Branches with >60 bootstrap support are labeled, and retrieved OTUs are depicted in bold. The closest environmental clones and cultivable representatives are also depicted with the corresponding GenBank accession number. Scale bar represents 10% estimated sequence divergence.

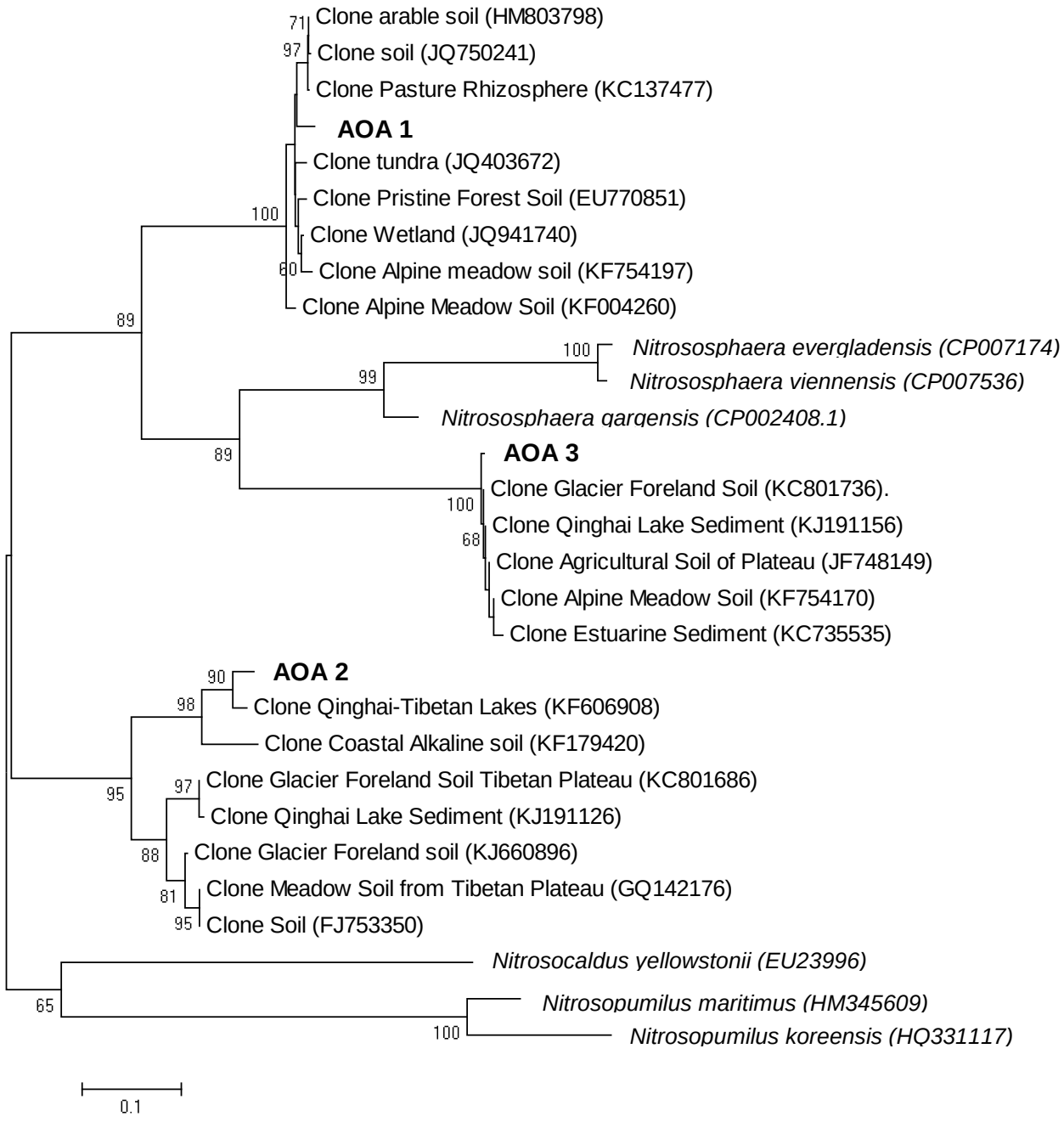


Figure 21.Unrooted *amoA*-based phylogenetic tree of ammonia-oxidizing Archaea with an 85% clustering cut-off. The 630-bp gene fragment using the archaeal primer set proposed by Francis et al. (2005) was used for phylogeny inference which was done in MEGA 6.0 (Tamura et al. 2013), using the NJ treeing method and the T92 + I + G nucleotide substitution model, with 1000 replicates to provide bootstrap support. Branches with >60 bootstrap support are labeled, and retrieved OTUs are depicted in bold. The closest environmental clones and cultivable representatives are also depicted with the corresponding GenBank accession number. Scale bar represents 10% estimated sequence divergence.

2. Bacterial *amoA*

For AOB, the mismatches between nucleotidic profiles and RFLP phylotypes were even more notorious: 222 clones were produced in the cloning step, corresponding to 5 distinct RFLP phylotypes (data not shown). Representatives of these phylotypes (n=73) were sequenced and clustered using Mothur using a cut-off of 80% and 96%. This yielded 2 and 6 OTUs respectively, and when compared with the RFLP phylotypes, there was no correspondence. Therefore, for AOB the initial RFLP profiling was unsuccessful and all further inferences were made using solely sequencing data.

Regarding AOB, the Darwin Mountains were the location with highest OTU diversity and frequency of occurrence at both clustering cut-offs with five out of six OTUs present with a clustering cut-off of 96% and two out of three OTUs present with a clustering cut-off of 80% (Fig. 22). The remaining locations have between one and two OTUs, with the highest

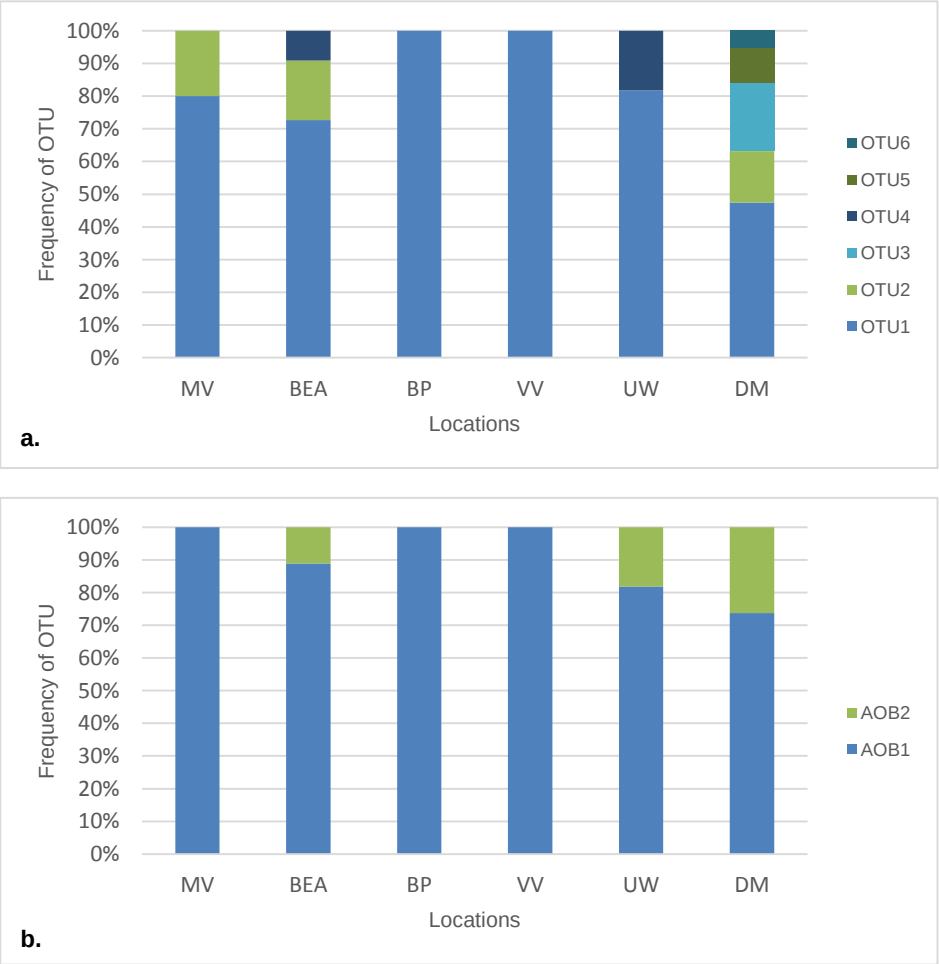


Figure 22. Number of retrieved bacterial OTUs with different clustering cut-offs, per location. a) a 96 % clustering cut-off yielded 6 bacterial OTUs, and b) a 80% clustering cut-off yielded 2 bacterial OTUs. MV- Miers Valley; BEA – Beacon Valley; BP- Battleship Promontory; VV – Victoria Valley; UW - Upper Wright Valley; DM – Darwin Mountains

clustering cut-off, and only one with the lowest clustering cut-off. With a 96% clustering cut-off, OTU1 was the most abundant containing 78% (n=57) of the clones, whereas the remaining 22% were distributed along OTU2 (n=6), OTU3 (n=4), OTU4 (n=3), OTU5 (n=2) and OTU6 (n=1). OTUs showed respectively 93-96%, 95-96%, 99%, 91-94%, 95-96%, 98-99% nucleotide sequence similarity to *amoA* gene sequences recovered from other environments such as supraglacial cryoconite (Cameron et al. 2012), alpine meadow soils (Xie & Ma, unpublished), and estuarine sediments (Chen et al. in press; Urakawa et al. unpublished). The placement of AOB OTUs in the NJ tree falls into two clusters each of them containing three OTUs (Fig. 23). OTU1, OTU2 and OTU5 seem to be closely affiliated with the *Nitrosospira* cluster, and appear to be sister group of nitrosospiras from cluster 3. OTU3, OTU4 and OTU6 have closer affiliations with environmental clones of estuarine ecosystems and salt marsh sediments, and also closer affiliations with members of the genus *Nitrosomonas*, namely *N. cryotolerans*, and the *N. marina* lineage. With an 80% clustering cut-off, two OTUs were obtained: AOB1 resulting from the merge of OTU1, OTU2, and OTU5; and AOB2 resulting from the merging of OTU3, OTU4 and OTU6 (Fig. 22b; Fig. 24). AOB1 was the most abundant OTU, with 89% of the clones (n=65), whereas the remaining 11% belonged to AOB2 (n=8). The tendencies and affiliations in the tree remain the same, as the ones observed with a higher clustering cut-off.

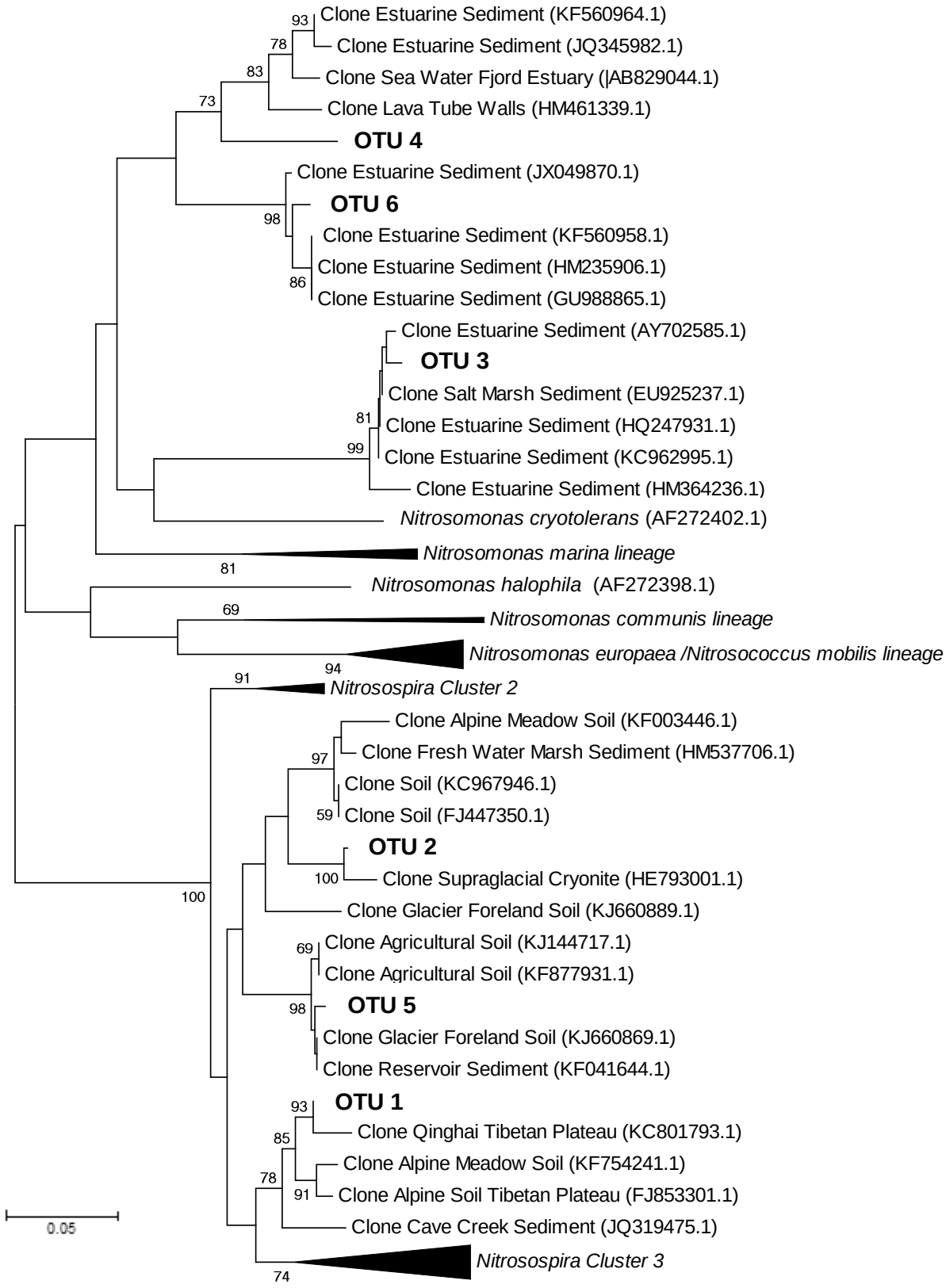


Figure 23. Unrooted *amoA*-based phylogenetic tree of ammonia-oxidizing Bacteria with a 96% clustering cut-off. The 490-bp gene fragment using the bacterial primer set proposed by Rotthauwe et al., (1997), and Okano et al., (2004) was used for phylogeny inference which was done in MEGA 6.0 (Tamura et al. 2013) using the NJ treeing method and the K92+ G nucleotide substitution model, with 1000 replicates to provide bootstrap support. Branches with >60 bootstrap support are labeled, and retrieved OTUs are depicted in bold. The closest environmental clones and cultivable representatives are also depicted with the corresponding GenBank accession number. Scale bar represents 5 % estimated sequence divergence.

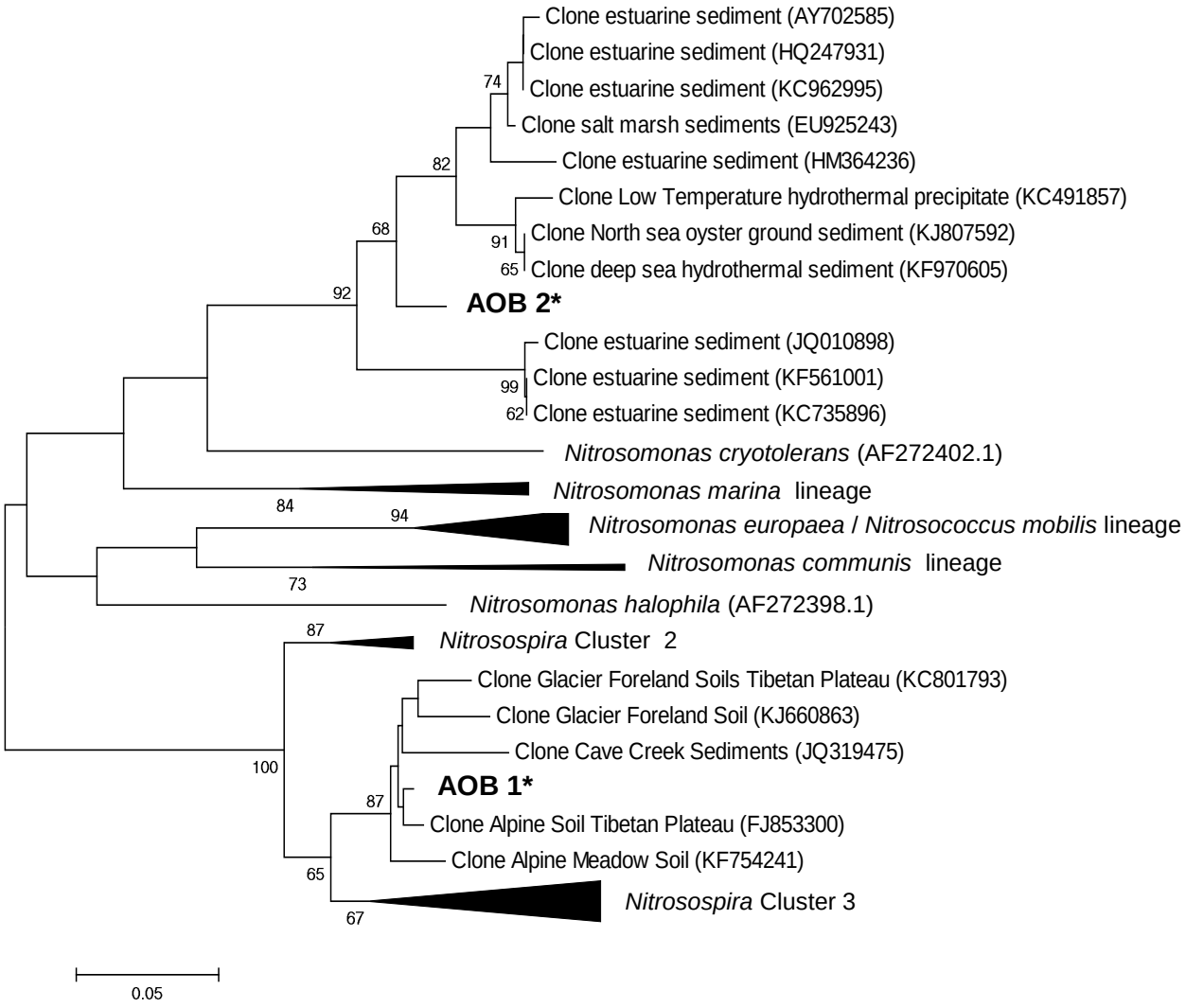


Figure 24.Unrooted *amoA*-based phylogenetic tree of ammonia-oxidizing Bacteria with a 80% clustering cut-off. The 490-bp gene fragment using the bacterial primer set proposed by Rotthauwe et al., (1997), and Okano et al., (2004) was used for phylogeny inference which was done in MEGA 6.0 (Tamura et al. 2013) using the NJ treeing method and the K92+ G nucleotide substitution model, with 1000 replicates to provide bootstrap support. Branches with >60 bootstrap support are labeled, and retrieved OTUs are depicted in bold. The closest environmental clones and cultivable representatives are also depicted with the corresponding GenBank accession number. Scale bar represents 5 % estimated sequence divergence.

V. Discussion

Biogeochemical processes in Antarctic ecosystems are limited by a set of extreme conditions, including low temperatures, moisture, nutrient availability, multiple freeze-thaw cycles, and high salinity stress (Bockheim 1997; Doran et al. 2002; Wall 2005; Aislabie et al. 2006), resulting in low microorganism diversity and abundance, which encompasses those involved in key geochemical processes. Nevertheless, it is important to understand system dynamics in these extreme environments, because ecosystem responses to climate variability are amplified in cold and high latitude regions and may provide early evidence of what could happen in the future, in milder ecosystems (Barrett et al. 2006b). Also, the existing limited diversity that seems to be ruled only by abiotic factors facilitates direct assessments of the contribution of particular species/genes to ecosystem processes (Cary et al. 2010; Richter et al. 2014).

Although Bacteria have been well documented within Antarctic soils (Aislabie et al. 2006), little is known about the presence or community composition of members of the Archaea domain in these environments, contrasting with the numerous 16S rRNA gene surveys performed in temperate environments, which have revealed the wide distribution of this group of microorganisms (Francis et al, 2005; Nicol & Schleper, 2006; Schleper & Nicol, 2010). General archaeal 16S rRNA gene sequences were only recently reported in the Ross Sea region (Ayton et al. 2010) and in the McMurdo Dry Valleys (Richter et al. 2014). In both locations, there was low diversity and dominance of certain OTUs, such as those belonging to the phylum Thaumarchaeota (99% and 80% of all archaeal 16S rRNA gene sequences, respectively in the Ross Sea and McMurdo Dry Valleys). These studies suggested that the majority of the archaeal representatives in Antarctic ecosystems and particularly in the Dry Valleys are involved in the nitrogen cycle and therefore possess may the *amoA* gene. A very recent quantitative study using the *amoA* gene as functional marker also revealed that AOA are prevalent in the Dry Valleys, dominating over AOB in some of them (Magalhães et al. 2014), as it has been observed in other environments (Leininger et al. 2006).

The majority of the archaeal clones generated in this study had strong affiliations with the *Nitrososphaera* cluster. Members of this group are most abundant and ubiquitous in soils and have been recovered from agricultural soils (Bintrim 1997), grassland soil, permafrost affected soils and soft limestone soil (Ochsenreiter et al. 2003), developing soils at deglaciated sites (Nicol & Schleper 2006) and barley field soils (Poplawski et al. 2007). Also, they are the only members of Thaumarchaeota detected in low nutrient grassland soil and in foreland of a receding glacier (Ochsenreiter et al. 2002; Nicol et al. 2005), which suggests that this lineage

is particularly adept of colonizing low nutrient soils devoid of vegetation. Particularly, *Nitrososphaera gargensis* has been shown to be inhibited by total ammonium concentrations in the lower mM range (Hatzenpichler et al. 2008), giving physiological support for a preference of low substrate concentrations by members of this lineage (Leininger et al. 2006). Comparatively to the *Nitrosopumilus* cluster, members of the *Nitrososphaera* cluster are also more resistant to freeze-thaw cycles, which can also explain their presence in the Transantarctic Mountains (Pesaro et al. 2003).

Using archaeal *amoA* gene sequences derived from pyrosequencing, Pester et al. (2012) have proposed that archaeal *amoA* gene sequences with less than 85% nucleic acid sequence identity are likely to represent two different AOA species. This threshold was calculated based on the assumption proposed by Stackenbrandt and Ebers, (2006) that a 98.5 % sequence identity at the 16S rRNA gene level was the approximate threshold below which microbes could be assigned to different species, thus providing the required DNA-DNA hybridization threshold of approximately 70% (Stackebrandt & Ebers, 2006; Tindall et al. 2010; Stackebrandt 2011). In their paper, Pester et al. (2012) obtained 83 AOA species-level OTUs based on their 85% cut-off, across a range of 16 soils, reaching the conclusion that the *Nitrososphaera* cluster dominated all soils and was probably composed of many still undescribed species (Pester et al. 2012).

Many ecological studies often use higher clustering cut-offs, especially when dealing with a functional gene like *amoA*, in order to have higher accuracy in detecting changes in gene function as well as different strains of a given species, across ecosystems characterized by distinct ecological features (Bernhard et al. 2010; Spang et al. 2010; Herrmann et al. 2012; Cao et al. 2013). When using a higher clustering cut-off (95%), we obtained five OTUs, three of them sharing less than 93% sequence identity with environmental clones and 77-80% sequence identity with members of the *Nitrososphaera* cluster, hinting to the fact that they could represent different species or represent highly adapted strains of one species (likely belonging to the *Nitrososphaera* genus). When adopting a lower clustering cut-off (85% proposed by Pester et al. (2012) the tendencies and affiliations were the same, suggesting that there are probably new species of AOA in these extreme environments affiliated with the *Nitrosophaera* cluster, which is plausible given the fact that AOA are still poorly studied and were only recently pointed as major contributors to the overall nitrogen budget (Könneke et al. 2005; Walker et al. 2010; Spang et al. 2012; Radax et al. 2012). In addition, there are still very few reports of ammonia-oxidizing microorganisms in the extreme Antarctic ecosystems and even less in the Dry Valleys (Magalhães et al. 2014; Richter et al. 2014).

The role of ammonia-oxidizing bacteria in nitrification was firstly described by Winogradsky (1890) and these microorganisms have been the target of numerous ecological studies for decades (Kowalchuk & Stephen 2001; Monteiro et al. 2014). They are placed among two distinct classes within the Proteobacteria phylum. Betaproteobacteria includes genera *Nitrosomonas* and *Nitrosospira*, which form a monophyletic clade, whereas AOB from Gammaproteobacteria are represented by a single polyphyletic genus, *Nitrosococcus* (Head et al. 1993; Purkhold et al. 2000; Purkhold et al. 2003). AOB from the Betaproteobacteria class are ubiquitous in soil and marine environments, but AOB from Gammaproteobacteria are exclusive of marine environments, with the only exception of the hyper saline Lake Bonney (Taylor Valley), where the existence members of the *Nitrosococcus* genus was reported in a 16S rRNA gene survey (Voytek et al.1999). Some lakes in the Dry Valleys, namely Lake Bonney have ion signatures that indicate a marine source (Green et al. 1988; Chinn 1993) and it has been suggested that they may have been formed when the area lifted isostatically from the sea and sea water was trapped in basins formed by large glaciers (Chinn 1993; Voytek et al. 1999).

Most of the ecological studies of AOB are focused in members of the Betaproteobacteria classe because this class is monophyletic and relatively accessible to molecular ecology analysis (Kowalchuk & Stephen 2001). In this study, the primer set used to amplify the bacterial *amoA* gene sequence was produced by Rotthauwe et al. (1997) and Okano et al. (2004) and is specific for members of the Betaproteobacteria class, therefore no inferences regarding the existence of members of *Nitrosococcus* genus were made.

Within AOB of the Betaproteobacteria class, advances in DNA-based techniques for direct microbial community analysis have revealed that *Nitrosospira* rather than *Nitrosomonas* are ubiquitous in many environments and are dominant in soil (Koops & Pommerening-Rose, 2001). Cultivable nitrosomonads can be subdivided into six lineages (using 16s rRNA and *amoA*) comprising eleven species, which are consistently retrieved using different treeing methods and which - if consisting of more than one sequence - have parsimony bootstrap values of above 90% (Koops & Purkhold 2006). Regarding nitrosospiras, although 16S rRNA phylogeny does not reveal an obvious substructure, Stephen et al (1996) suggested a subdivision into 4 clusters: clusters 2, 3 and 4 containing culturable species, and cluster 1 containing uncultivable representatives. Purkhold et al. (2000) expanded this system by proposing the addition of cluster 0, composed of strains retrieved from undisturbed/unfertilized soils (Purkhold et al. 2000; Nugroho et al. 2005). These clusters are generally found with all treeing methods, but not all of them are well supported by bootstrap analysis. Other phylogenetic markers have been tested for phylogenetic inference of AOB, namely the

intergenic region between 16S and 23s rRNA (Aakra et al. 2001b), but due to inconsistencies between the 16S rRNA and the ITS tree topologies, and the fact that the latter marker is highly variable leading to difficulties in alignments, inferences using ITS should be carefully analyzed (Koops & Purkhold 2006). The *amoA* gene has also been widely used as an additional marker (McTavish et al. 1993; Rotthauwe et al. 1997; Aakra et al. 2001a; Purkhold et al. 2000; Purkhold et al. 2003) and tree topologies show high consistency with previously 16S rRNA-based phylogeny. In their paper, Koops et al. (2006) used a generally used primer set which amplified 453-bp of the *amoA* gene and concluded that, if compared with the 16S rRNA, amoA analysis provides less resolution, given the fact that a relatively short (151 aminoacids) and highly conserved aminoacid sequence stretch is used as a marker. This can be overcome by amplifying longer fragments of the *amoA* gene, which has already been attempted (Norton et al. 2002).

Similarly to Pester et al. (2012), Koops et al (2006) generated correlation plots of bacterial *amoA*/amoA similarity versus 16S rRNA similarity of all possible pairs of Betaproteobacteria isolates available at the time and demonstrated, among other things, that AOB showing less than 80% nucleic acid similarity (or 85% aminoacidic similarity) always possess less than the currently accepted 16S rRNA threshold value for bacterial genospecies (Stackebrandt & Goebel 1994). Hence, *amoA* nucleic sequences of a newly AOB isolate with a lower similarity percentage than those thresholds would be indicative of a previously undescribed species (Koops & Purkhold 2006).

Similarly to what is seen in ecological studies focused in AOA, general clustering cut-offs are often higher than the one suggested by Koops et al. (2006) and our results using a similar cut-off (96%) yielded an attribution of 6 OTUs divided into distinct clusters containing members of the *Nitrosomonas* and *Nitrosospira* genera (Fig. 23; Fig. 24). Clusters 2 and 3 of *Nitrosospira* are depicted, and OTUs seem to be closely affiliated with cluster 3, which is also the most ubiquitous in soils with neutral pH (Kowalchuk & Stephen 2001; Nugroho et al. 2005). The other OTUs have stronger affiliations with environmental clones from estuarine ecosystems, and genus *Nitrosomonas,* namely *N. cryotolerans* and the *N. marina* cluster (which also contains *N. aestuarii*). Interestingly, these species of *Nitrosomonas* are known to be strictly halophilic (Koops & Pommerening-Roser 2001; Bernhard & Bollmann 2010) and their association with our OTUs, even if not well supported, may indicate an affinity for saline locations. When looking at the tree obtained with a lower clustering cut-off (Fig. 24), AOB2 probably corresponds to an undescribed species of AOB, affiliated with *N. cryotolerans.* This particular species was isolated from a coastal bay in Alaska in 1988 (Jones & Morita 1988) and constitutes a single lineage among the nitrosomonads. Efforts towards genome

sequencing are already being performed (Nordberg et al. 2014), which will provide valuable insights regarding phylogenomics and niche adaptation of this species of AOB.

The Dry Valleys are divided in three microclimate zones, which are defined based on environmental parameters such as temperature, moisture, conductivity, geomorphology and lithology : a freeze-thaw zone (comprising Miers, Garwood and Marshal valleys), an inland mixed zone (including Victoria Valley, Battleship Promontory and Bull Pass), and an inland zone (including Beacon and Upper Wright valleys) (Lee et al. 2012). Miers Valley is the richest location in terms of archaeal *amoA* OTUs and also the one encompassing three unique archaeal OTUs. This valley is considered to have milder environmental characteristics and has been the scenario of many microbiological surveys, namely regarding Cyanobacteria (Wall 2005; Sokol et al. 2013; Yung et al. 2014), which are considered to represent the major source of nitrogen fixation in the Dry Valleys. It is an eastern, low altitude coastal valley, with the highest marine influence, highest pH, and the one the highest C/N ratio of all sampling sites, probably explained by the presence of a large central lake (Cowan 2009; Magalhães et al. 2012; Lee et al. 2012).

The Darwin Mountains are pointed as the richest location in terms of AOB *amoA* OTU diversity and abundance, with three unique OTUs, one obtained from the proximities of Lake Wellman, and two others from Junction Spur. Regarding archaeal OTUs it is the second richest location. Microbiological studies in these locations are still emerging, namely using community fingerprinting methods (Webster-Brown et al. 2010; Aislabie et al. 2011), revealing that factors such as terrain age have particular influence in the existing bacterial diversity (Magalhães et al. 2012; Aislabie et al. 2013). Comparatively to the Dry Valleys, the Darwin Mountains present milder environmental conditions, namely water availability and organic carbon derived from cianobacterial mats (Barrett et al. 2006b; Webster-Brown et al. 2010), potentially harboring a greater amount of life forms and more complex trophic webs. However, and even though we performed sequence quality checks, there are OTUs that are represented by only one or two sequences. Therefore, additional sampling and sequencing effort should be performed in the future, to draw more robust conclusions regarding the uniqueness of these OTUs.

Our study confirms the existence of ammonia-oxidizing microorganisms in the Dry Valleys, and provides the first evidence of AOA and AOB in the Darwin Mountains and in the Victoria Valley, pointing to the fact that the Darwin Mountains may harbor distinct lineages from those existing in other Dry Valleys. In both groups of ammonia-oxidizing microorganisms, there seems to exist cosmopolitan OTUs, with close affiliations with environmental clones from a

broad range of environments, suggesting that the functionality inferred by *amoA* gene sequences is maintained.

VII. Conclusions

The recent discovery of a new player in the ammonia-oxidation pathway of nitrification – AOA - demanded a reassessment of the relative roles of each group of ammonia-oxidizers, across a wide range of environments. The advent of new sequencing technologies is also providing valuable insights regarding the physiology, functionality and niche adaptation of ammonia-oxidizing microorganisms by providing direct access of full genomes isolated from different environments.

This study complemented the one of Magalhães et al. (2014) by adding two more locations across the Transantarctic Mountains (Darwin Mountains and Victoria Valley) and by reporting for the first time the existence of AOA and AOB in these locations. Regarding both groups of microorganisms, the retrieved OTUs showed higher affiliations with culturable members known to inhabit soil, which makes sense since the sampling locations are classified as hyper-arid polar deserts, in which water is scarce. As shown with two different clustering cut-offs, there, for both groups of ammonia-oxidizing microorganisms, there is a prevalence of cosmopolitan OTUs (that show affiliations with environmental clones from a wide range of environments), contrasting with OTUs that exist in a very low frequency and whose environmental affiliations are restricted. This shows that functionality seems to be maintained across the Dry Valleys ecosystem, and that there may exist few highly adapted still undescribed species or strains, in these extreme environments.

Future work will include quantitative assessments of the AOA and AOB *amoA* gene in all locations, in order to have a full characterization of both abundance and diversity of ammonia-oxidizing microorganisms in these extreme environments. In addition, it would also be interesting to check for differences regarding the aminoacidic sequences in the whole *amoA*, given the fact that it encodes the active site of the AMO enzyme and changes in amoA would have a direct effect on the whole enzyme. Particularly in extremely cold environments as the Transantarctic Mountains, it would be interesting to observe whether there are changes in protein structure, derived from mechanisms of cold adaptation. To achieve this, we would need to design primers that amplify a bigger fragment of the *amoA* gene sequence (and preferably the whole fragment) in both groups of ammonia-oxidizing microorganisms.

Final remarks and future perspectives

Antarctica is seen as the most pristine location on Earth, and a baseline against which to monitor future global changes. Its biodiversity is massively undersampled and may be a source of still undiscovered genetic resources with applications in almost every field of Science.

Culture-independent methods have provided great insights in the field of Microbial Ecology. In this study, we evaluated: 1) the existing bacterial richness and diversity in one recently sampled valley, using community fingerprinting methods; 2) the effect of one particular environmental parameter (water availability) in bacterial community structure across a transect with increasing water distance. A high throughput sequencing approach was used, using the 16S rRNA gene as a molecular marker; and 3) the diversity and abundance of a particular group of microorganisms with a key role in the N cycle, by tagging a functional gene and performing an inter-valley comparison.

Results emphasize that the high degree of physicochemical heterogeneity and the extreme conditions which characterize the ice-free regions of Antarctica, clearly have a direct effect in shaping the distribution and the functional attributes of the existing microbial communities.

Future working perspectives will include new sampling expeditions, new experimental designs, and a cross-check of data between the multidisciplinary teams that constitutes the NZTABS/ICTAR International Programs. Particularly with the advent of new sequencing technologies and with the decrease of the sequence cost per base pair, it will be possible to go beyond 16S rRNA gene surveys and encompass full genomes, which will provide clearer insights regarding not only the genetic diversity, but also functionality of microorganisms inhabiting Antarctic ecosystems.

References

Aakra A, Utåker JB, and Nes IF. 2001a. “Comparative Phylogeny of the Ammonia Monooxygenase Subunit A and 16S rRNA Genes of Ammonia-Oxidizing Bacteria.” *FEMS Microbiology Letters* 205 (2): 237–242.

Aakra A. 2001b. “Detailed Phylogeny of Ammonia-Oxidizing Bacteria Determined by rDNA Sequences and DNA Homology Values. *International Journal of Systematic and Evolutionary Microbiology* 201: 2021–2030.

Abràmoff, MD. 2004. “Image Processing with ImageJ.” *Biophotonics International* 11 (7): 36-42

Aislabie JM, Bockheim J, Mcleod M, Hunter D, Stevenson B and. Barker GM. 2011. “Microbial Biomass and Community Structure Changes along a Soil Development Chronosequence near Lake Wellman, Southern Victoria Land.” *Antarctic Science* 24 (2): 154–164.

Aislabie JM, Lau A, Dsouza M, Shepherd C, Rhodes P, and Turner SJ. 2013. “Bacterial Composition of Soils of the Lake Wellman Area, Darwin Mountains, Antarctica.” *Extremophiles : Life under Extreme Conditions* 17 (5): 775-786

Aislabie JM, Chhour K-L, Saul DJ, Miyauchi S, Ayton J, Paetzold RF, and Balks MR. 2006. “Dominant Bacteria in Soils of Marble Point and Wright Valley, Victoria Land, Antarctica.” *Soil Biology and Biochemistry* 38 (10): 3041–3056.

Arp DJ, Chain PSG, and Klotz MG. 2007. “The Impact of Genome Analyses on Our Understanding of Ammonia-Oxidizing Bacteria.” *Annual Review of Microbiology* 61: 503–528.

Arp, DJ, Sayavedra-Soto LA, and Hommes NG. 2002. “Molecular Biology and Biochemistry of Ammonia Oxidation by Nitrosomonas Europaea.” *Archives of Microbiology* 178 (4): 250–255.

Ayton J, Aislabie JM, and Barker GM. 2010. “Crenarchaeota Affiliated with Group 1.1 B Are Prevalent in Coastal Mineral Soils of the Ross Sea Region of Antarctica.” *Environmental Microbiology* 12 (3): 689-670.

Babalola OO, Bronwyn M K, Roes-Hill M, Cook AE, Cary SC, Burton SG, and Cowan DA. 2009. “Phylogenetic Analysis of Actinobacterial Populations Associated with Antarctic Dry Valley Mineral Soils.” *Environmental Microbiology* 11 (3): 566–576.

Baker BJ, Sheik CS, Taylor CA, Jain S, Bhasi A, Cavalcoli JD, and Dick GJ. 2013. “Community Transcriptomic Assembly Reveals Microbes That Contribute to Deep-Sea Carbon and Nitrogen Cycling.” *The ISME J.* 7 (10):1962–1973.

Barrett JE, Virginia RA, Wall DH, Cary SC, Adams BJ, Hacker AL, and Aislabie JM. 2006a. “Co-Variation in Soil Biodiversity and Biogeochemistry in Northern and Southern Victoria Land, Antarctica.” *Antarctic Science* 18 (04): 535.

Barrett JE, Virginia RA, Hopkins DW, Aislabie JM, Bargagli R, Bockheim JG, Campbell IB, Lyons WF, Moorhead DL, Nkem JN, Sletten RS, Steltzer H, Wall DH and Wallenstein MD. 2006b. “Terrestrial Ecosystem Processes of Victoria Land, Antarctica.” *Soil Biology and Biochemistry* 38 (10): 3019–3034.

Bernhard A, Zachary E, Landry C, Blevins A, de la Torre JR, Giblin AE, and Stahl DA. 2010. "Abundance of Ammonia-Oxidizing Archaea and Bacteria along an Estuarine Salinity Gradient in Relation to Potential Nitrification Rates." *Applied and Environmental Microbiology* 76 (4): 1285–1289.

Bernhard AE and Bollmann A. 2010. "Estuarine Nitrifiers: New Players, Patterns and Processes." *Estuarine, Coastal and Shelf Science* 88 (1). Elsevier Ltd: 1–11.

Bintrim SB. 1997. "Molecular Phylogeny of Archaea from Soil." *PNAS* 94 (1): 277–282.

Bockheim JG and McLeod M. 2013. "Glacial Geomorphology of the Victoria Valley System, Ross Sea Region, Antarctica." *Geomorphology* 193:14–24.

Bockheim, JG. 1997. "Properties and Classification of Cold Desert Soils from Antarctica." *Soil Science Society of America Journal* 61:224-231.

Bolhuis H, Palm P, Wende A, Falb M, Rampp M, Rodriguez-Valera F, Pfeiffer F, and Oesterhelt D. 2006. "The Genome of the Square Archaeon Haloquadratum Walsbyi : Life at the Limits of Water Activity." *BMC Genomics* 7 (January): 169-181.

Boon N, Windt W, Verstraete W and Top EM. 2002. "Evaluation of Nested PCR-DGGE (denaturing Gradient Gel Electrophoresis) with Group-Specific 16S rRNA Primers for the Analysis of Bacterial Communities from Different Wastewater Treatment Plants." *FEMS Microbiology Ecology* 39 (2):101–112.

Bottos EM, Scarrow JW, Archer SDJ, Mcdonald IR, and Cary SC. 2014. "Bacterial Community Structure in Antarctic Soils." In *Antarctic Terrestrial Microbiology* Chap. 2, Don A. Cowan Ed.: 9–34. Berlin, Heidelberg: Springer, Berlin Heidelberg.

Brambilla E, Hippe H, Hagelstein A, Tindall BJ, Stackebrandt E. 2001. "16S rDNA Diversity of Cultured and Uncultured Prokaryotes of a Mat Sample from Lake Fryxell, McMurdo Dry Valleys, Antarctica." *Extremophiles* 5: 23–33.

Brochier-Armanet C, Boussau B, Gribaldo S and Forterre P. 2008. "Mesophilic Crenarchaeota: Proposal for a Third Archaeal Phylum, the Thaumarchaeota." *Nature Reviews. Microbiology* 6 (3): 245–252.

Brochier-Armanet C, Forterre P, and Gribaldo S. 2011. "Phylogeny and Evolution of the Archaea: One Hundred Genomes Later." *Current Opinion in Microbiology* 14 (3): 274–281.

Burke C, Steinberg P, Rusch D, Kjelleberg S, and Thomas T. 2011. "Bacterial Community Assembly Based on Functional Genes rather than Species." *PNAS* 108 (34): 14288–14293.

Calkin PE. 1971. "Glacial Geology of the Victoria Valley System, Southern Victoria Land, Antarctica." In *Antarctic Snow and Ice Studies II*, (A.P. Crary Ed.): 363–412. American Geophysical Union, Washington, D. C.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and Madden TL. 2009. "BLAST+: Architecture and Applications." *BMC Bioinformatics* 10: 421-429

Cameron KA, Hodson AJ, and Osborn AM. 2012. "Carbon and Nitrogen Biogeochemical Cycling Potentials of Supraglacial Cryoconite Communities." *Polar Biology* 35 (9): 1375–1393.

Cameron RE. 1972. "Microbial and Ecologic Investigations in Victoria Valley, Southern Victoria Land, Antarctica." *Antarctic Research Series* 20: 195-260. American Geophysical Union, Washington, D. C.

Cao H, Auguet JC, and Gu J-D. 2013. “Global Ecological Pattern of Ammonia-Oxidizing Archaea.” *PloS One* 8 (2): e52853.

Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, and Knight R. 2010a. “PyNAST: A Flexible Tool for Aligning Sequences to a Template Alignment.” *Bioinformatics* 26 (2): 266–267.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, D Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koening JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaeveld J and Knight R. 2010b. “QIIME Allows Analysis of High-Throughput Community Sequencing Data.” *Nature Methods* 7 (5): 335–336.

Cary SC, McDonald IR, Barrett JE, and Cowan DA. 2010. “On the Rocks: The Microbiology of Antarctic Dry Valley Soils.” *Nature Reviews. Microbiology* 8 (2): 129–138.

Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, and Kjelleberg S. 2007. “Use of 16S rRNA and rpoB Genes as Molecular Markers for Microbial Ecology Studies.” *Applied and Environmental Microbiology* 73 (1): 278–288.

Cavicchioli R, and Siddiqui KS. 2002. “Low-Temperature Extremophiles and Their Applications.” *Current Opinion in Biotechnology* 13(3): 253–261.

Chan JZ-M, Halachev MR, Loman NJ, Constantinidou C, and Pallen MJ. 2012. “Defining Bacterial Species in the Genomic Era: Insights from the Genus *Acinetobacter*.” *BMC Microbiology* 12 (1). BMC Microbiology: 302.

Chan Y, Van Nostrand JD, Zhou J, Pointing SB, and Farrell RL. 2013. “Functional Ecology of an Antarctic Dry Valley.” *PNAS* 110 (22): 8990–8995.

Chanal A, Chapon V, Benzerara K, Barakat M, Christen R, Achouak W, Barras F and Heulin T. 2006. “The Desert of Tataouine: An Extreme Environment That Hosts a Wide Diversity of Microorganisms and Radiotolerant Bacteria.” *Environmental Microbiology* 8 (3). Blackwell Publishing Ltd: 514–25.

Chinn, TJ. 1993. Physical Hydrology of the Dry Valley Lakes, in *Physical and Biogeochemical Processes in Antarctic Lakes* (eds W. J. Green and E. I. Friedmann), American Geophysical Union, Washington, D. C.. doi: 10.1029/AR059p0001

Chotte, JL. 2005. “Importance of Microorganisms for Soil Aggregation.” in *Microorganisms in Soils: Roles in Genesis and Functions*, Volume 3, 2005: 107-119 - Springer Berlin Heidelberg

Chu, H, Fierer N, Lauber NC, Caporaso JG, Knight R, and Grogan P. 2010. “Soil Bacterial Diversity in the Arctic Is Not Fundamentally Different from That Found in Other Biomes.” *Environmental Microbiology* 12 (11): 2998–3006.

Chwedorzewska K. 2009. “Terrestrial Antarctic Ecosystems in the Changing World: An Overview.” *Polish Polar Research* 30 (3): 263–76.

Clarke, A, Barnes DKA, and Hodgson DA. 2005. “How Isolated Is Antarctica?” *Trends in Ecology & Evolution* 20 (1): 1–3.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM and Tiedje JM. 2009. “The Ribosomal Database Project: Improved Alignments and New Tools for rRNA Analysis.” *Nucleic Acids Research* 37 (Database issue): D141–D145.

Cole JR and Tiedje JM. 2014. “History and Impact of RDP.” *RNA Biology* 11 (3): 239–243.

Convey P, Gibson JAE, Hillenbrand CD, Hodgson DA, Philip Pugh PJ, Smellie JL and Stevens MI. 2008. “Antarctic Terrestrial Life--Challenging the History of the Frozen Continent?” *Biological Reviews of the Cambridge Philosophical Society* 83 (2): 103–17.

Cordero OX, and Polz MF. 2014. “Explaining Microbial Genomic Diversity in Light of Evolutionary Ecology.” *Nature Reviews. Microbiology* 12 (4). Nature Publishing Group: 263–273.

Cowan DA. 2014. *Antarctic Terrestrial Microbiology*. *Springer 2014*

Cowan DA. 2009. “Cryptic Microbial Communities in Antarctic Deserts.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (47): 19749–19750.

Culman SW, Bukowski R, Gauch HG, Cadillo-Quiroz H, and Buckley DH. 2009. “T-REX: Software for the Processing and Analysis of T-RFLP Data.” *BMC Bioinformatics* 10 (January): 171.

Curtis TP, Sloan WT and Scannell JW. 2002. “Estimating Prokaryotic Diversity and Its Limits.” *Proceedings of the National Academy of Sciences of the United States of America* 99 (16): 10494–10499.

Daebeler A, Abell GC, Bodelier PL, Bodrossy L, Frampton DM, Hefting MM and Laanbroek HJ. 2012a. “Archaeal Dominated Ammonia-Oxidizing Communities in Icelandic Grassland Soils Are Moderately Affected by Long-Term N Fertilization and Geothermal Heating.” *Frontiers in Microbiology* 3 (10): 1–14.

Dahllöf I, Baillie H, Kjelleberg S and Dahllo I. 2000. “rpoB -Based Microbial Community Analysis Avoids Limitations Inherent in 16S rRNA Gene Intraspecies Heterogeneity.” *Applied and Environmental Microbiology* 66 (8): 3376–3380.

Darling CA and Siple PA. 1940. “Bacteria of Antarctica.” *Journal of Bacteriology* 42 (a): 83–98.

De La Torre JR, Walker CB, Ingalls AE, Könneke M and Stahl DA. 2008. “Cultivation of a Thermophilic Ammonia Oxidizing Archaeon Synthesizing Crenarchaeol.” *Environmental Microbiology* 10 (3): 810–818.

Delsuc F, Brinkmann H and Philippe H. 2005. “Phylogenomics and the Reconstruction of the Tree of Life.” *Nature Reviews. Genetics* 6 (5): 361–375.

Deming JW. 2002. “Psychrophiles and Polar Regions.” *Current Opinion in Microbiology* 5 (3): 301–309.

DeSantis T Z, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P and Andersen GL. 2006. “Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB.” *Applied and Environmental Microbiology* 72 (7): 5069–5072.

Di Prisco G, Convey P, Gutt J, Cowan D, Conlan K and Verde C. 2012. “Understanding and Protecting the World’s Biodiversity: The Role and Legacy of the SCAR Programme ‘Evolution and Biodiversity in the Antarctic’.” *Marine Genomics* 8 (December). Elsevier B.V. 3–8.

Doran PT, Priscu JC, Lyons WB, Walsh JE, Fountain AG, McKnight DM, Moorhead DL, Virginia AG, Wall DH, Clow GD, Fritsen CH, McKay CP and Parsons AN. 2002. “Antarctic Climate Cooling and Terrestrial Ecosystem Response.” *Nature* 415: 517–520.

Edgar R C. 2010. “Search and Clustering Orders of Magnitude Faster than BLAST.” *Bioinformatics (Oxford, England)* 26 (19): 2460–2461.

Edgar RC, Haas BJ, Clemente JC, Quince C and Knight R. 2011. “UCHIME Improves Sensitivity and Speed of Chimera Detection.” *Bioinformatics* 27 (16): 2194–2200.

Edmeades DC, Wheeler DM and Clinton OE. 1985. “The Chemical Composition and Ionic Strength of Soil Solutions from New Zealand Topsoils.” *Soil Research* 23 (2): 151–165.

Falkowski PG, Fenchel T, and Delong EF. 2008. “The Microbial Engines That Drive Earth’s Biogeochemical Cycles.” *Science* 320 (5879): 1034–1039.

Farrelly V1995. “Effect of Genome Size and Rrn Gene Copy Number on PCR Amplification of 16S rRNA Genes from a Mixture of Bacterial Species.” *Applied and Environmental Microbiology* 61 (7): 2798-2801

Feller G and Gerday C. 2003. “Psychrophilic Enzymes: Hot Topics in Cold Adaptation.” *Nature Reviews. Microbiology* 1 (3): 200–208.

Felsenstein J. 1989. “PHYLIP - Phylogeny Inference Package Version 3.2.” *Cladistics* 5 (2). Blackwell Publishing Ltd: 163–166.

Fierer N, Breitbart M, Nulton J, Salamon P, Lozupone C, Jones R, Robeson M, Edwards RA, Felts B, Rayhawk S, Knight R, Rohwer F and Jackson RB. 2007. “Metagenomic and Small-Subunit rRNA Analyses Reveal the Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil.” *Applied and Environmental Microbiology* 73 (21): 7059–7066.

Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, Owens S, Gilbert JA, Wall DH and Caporaso JG. 2012. “Cross-Biome Metagenomic Analyses of Soil Microbial Communities and Their Functional Attributes.” *PNAS* 109 (52): 21390–21395.

Fisher MM and Triplett EW. 1999. “Automated Approach for Ribosomal Intergenic Spacer Analysis of Microbial Diversity and Its Application to Freshwater Bacterial Communities.” *Applied and Environmental Microbiology* 65 (10): 4630–4636.

Flint E and. Stout JD. 1960. “Microbiology of Some Soils from Antarctica.” *Nature* 188: 767–768.

Francis CA, Beman JM and Kuypers MMM. 2007. “New Processes and Players in the Nitrogen Cycle: The Microbial Ecology of Anaerobic and Archaeal Ammonia Oxidation.” *The ISME J.* 1 (1): 19–27.

Francis CA, Roberts KR, Beman JM, Santoro AE and Oakley BB. 2005. “Ubiquity and Diversity of Ammonia-Oxidizing Archaea in Water Columns and Sediments of the Ocean.” *PNAS* 102 (41): 14683–14688.

Fu L, Niu B, Zhu Z, Wu S and Li W. 2012. “CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data.” *Bioinformatics* 28 (23): 3150–3152.

Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W and Bryant SH. 2010. “The NCBI BioSystems Database.” *Nucleic Acids Research* 38 (Database issue): D492–D496.

Gilbert JA, Jansson JK and Knight R. 2014. “The Earth Microbiome Project: Successes and Aspirations.” *BMC Biology* 12 (1): 69.

Glaser K, Hackl E, Inselsbacher E, Strauss J, Wanek W, Zechmeister-Boltenstern S and Sessitsch A. 2010. “Dynamics of Ammonia-Oxidizing Communities in Barley-Planted Bulk Soil and Rhizosphere Following Nitrate and Ammonium Fertilizer Amendment.” *FEMS Microbiology Ecology* 74 (3): 575–591.

Grant WD. 2004. “Life at Low Water Activity.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 359 (1448): 1249–1267.

Green WJ, Angle MP and Chave KE. 1988. “The Geochemistry of Antarctic Streams and Their Role in the Evolution of Four Lakes of the McMurdo Dry Valleys.” *Geochimica et Cosmochimica Acta* 52 (5): 1265–1274.

Gribaldo S and Brochier C. 2009. “Phylogeny of Prokaryotes: Does It Exist and Why Should We Care?” *Research in Microbiology* 160 (7): 513–521.

Gupta P, Reddy GSN, Delille D and Shivaji S. 2004. “*Arthrobacter gangotriensis* sp. nov. and *Arthrobacter kerguelensis* sp. nov. from Antarctica.” *International Journal of Systematic and Evolutionary Microbiology* 54 (6): 2375–2378.

Hall BL, Dento GH, Overturf B and Hendy CH. 2002. “Glacial Lake Victoria, a High-Level Antarctic Lake Inferred from Lacustrine Deposits in Victoria Valley.” *Journal of Quaternary Science* 17 (7): 697–706.

Handelsman J. 2004. “Metagenomics: Application of Genomics to Uncultured Microorganisms.” *Microbiology and Molecular Biology Reviews* 68 (4): 669-685

Harley S. 1988. “Proterozoic Granultites from the Rauer Group, East Antarctica .1. Decompressional Pressure Temperature Paths Deducedd from Mafic and Felsic Gneisses.” *Journal of Petrology* 29 (5): 1059–1095.

Hatzenpichler R. 2012. “Diversity, Physiology, and Niche Differentiation of Ammonia-Oxidizing Archaea.” *Applied and Environmental Microbiology* 78 (21): 7501–7510.

Hatzenpichler R, Lebedeva EV, Spieck E, Stoecker K, Richter A, Daims H, and Wagner M. 2008. “A Moderately Thermophilic Ammonia-Oxidizing Crenarchaeote from a Hot Spring.” *PNAS* 105 (6): 2134–2139.

Hayakawa M and Nonomura H. 1987. “Humic Acid-Vitamin Agar, a New Medium for the Selective Isolation of Soil Actinomycetes.” *Journal of Fermentation Technology* 65 (5): 501–509.

Hayatsu M, Tago K and Saito M. 2008. “Various Players in the Nitrogen Cycle: Diversity and Functions of the Microorganisms Involved in Nitrification and Denitrification.” *Soil Science and Plant Nutrition* 54 (1): 33–45.

Head IM, Hiorns WD, Embley TM, McCarthy AJ and Saunders JR. 1993. “The Phylogeny of Autotrophic Ammonia-Oxidizing Bacteria as Determined by Analysis of 16S Ribosomal RNA Gene Sequences.” *Journal of General Microbiology* 139 (6): 1147–1153.

Herrmann M, Hädrich A and Küsel K. 2012. “Predominance of Thaumarchaeal Ammonia Oxidizer Abundance and Transcriptional Activity in an Acidic Fen.” *Environmental Microbiology* 14 (11): 3013–3025.

Hogg ID, Cary SC, Convey P, Newsham KK, O’Donnell AG, Adams BJ, Aislabie J, Frati F,. Stevens MI and Wall DH. 2006. “Biotic Interactions in Antarctic Terrestrial Ecosystems: Are They a Factor?” *Soil Biology and Biochemistry* 38 (10): 3035–3040.

Hopkins DW, Sparrow AD, Novis PM, Gregorich EG, Elberling B, and Greenfield LG. 2006. “Controls on the Distribution of Productivity and Organic Resources in Antarctic Dry Valley Soils.” *Proceedings. Biological Sciences / The Royal Society* 273 (1602): 2687–2695.

Hugenholtz P. 2002. “Exploring Prokaryotic Diversity in the Genomic Era.” *Genome Biology* 3 (2): :reviews0003.1–0003.8

Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA and Sogin ML. 2008. “Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing.” *PLoS Genetics* 4 (11): e1000255.

Janssen J and Prosser JI. 2013. “The Life beneath Our Feet.” *Nature* 494: 7–8.

Jia Z and Conrad R. 2009. “Bacteria rather than Archaea Dominate Microbial Ammonia Oxidation in an Agricultural Soil.” *Environmental Microbiology* 11 (7): 1658–1671.

Jonasson J, Olofsson M and Monstein H-J. 2002. “Classification, Identification and Subtyping of Bacteria Based on Pyrosequencing and Signature Matching of 16S rDNA Fragments.” *APMIS* 110 (3): 263–272.

Jones, RD and Morita RY. 1988. “A New Marine Ammonium-Oxidizing Bacterium, *Nitrosomonas cryotolerans* sp. nov.” *Canadian Journal of Microbiology*: 1022-1028

Jung J, Yeom J, Kim J, Han J, Lim HS, Park H, Hyun S and Park W. 2011. “Change in Gene Abundance in the Nitrogen Biogeochemical Cycle with Temperature and Nitrogen Addition in Antarctic Soils.” *Research in Microbiology* 162 (10): 1018–1026.

Jung M-Y, Well R, Min D, Giesemann A, Park S-J, Kim J-G, Kim S-J and Rhee S-K. 2014. “Isotopic Signatures of N2O Produced by Ammonia-Oxidizing Archaea from Soils.” *The ISME J.* 8 (5): 1115–1125.

Junier P, Molina V, Dorado C, Hadas O, Kim O-S, Junier T, Witzel JP and F Imhoff JF. 2010. “Phylogenetic and Functional Marker Genes to Study Ammonia-Oxidizing Microorganisms (AOM) in the Environment.” *Applied Microbiology and Biotechnology* 85 (3): 425–440.

Klindworth A , Pruesse E, Schweer T, Peplies J, Quast C, Horn M and Glöckner FO. 2013. “Evaluation of General 16S Ribosomal RNA Gene PCR Primers for Classical and next-Generation Sequencing-Based Diversity Studies.” *Nucleic Acids Research* 41 (1): 1-11

Könneke M, Bernhard AE, De La Torre JR, Walker CB, Waterbury JB and Stahl DA. 2005. “Isolation of an Autotrophic Ammonia-Oxidizing Marine Archaeon.” *Nature* 437: 543–546.

Konstantinidis K, T, Ramette A and Tiedje JM. 2006. “The Bacterial Species Definition in the Genomic Era.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 361:1929–1940.

Koops HP, and Pommerening-Röser A. 2001. “Distribution and Ecophysiology of the Nitrifying Bacteria Emphasizing Cultured Species.” *FEMS Microbiology* 37: 1–9.

Koops HP, Purkhold U, Pommerening-Röser A, Timmermann G, Wagner M 2006. “The Lithoautotrophic Ammonia-Oxidizing Bacteria.” in*The Prokaryotes*, M Dworkin, S Falkow, E Rosenberg, KH Schleicher, E Stackebrandt Ed, pg.778–811. Springer Verlag

Kowalchuk GA and Stephen JR. 2001. “Ammonia-Oxidizing Bacteria: A Model for Molecular Microbial Ecology.” *Annual Review of Microbiology* 55: 485–529.

Kowalchuk GA and Stephen JR. 1997. “Beta Subdivision of the Class Proteobacteria in Coastal Sand Dunes by Denaturing Gradient Gel Electrophoresis and Sequencing of PCR-Amplified 16S Ribosomal DNA.” *Applied and Environmental Microbiology* 63 (4): 1489-1497

Kozlowski JA, Price J and Stein LY. 2014. “Revision of N2O-Producing Pathways in the Ammonia-Oxidizing Bacterium Nitrosomonas Europaea ATCC 19718.” *Applied and Environmental Microbiology* 80 (16): 4930–4935.

Kuczynski J, Lauber CL, Walters WA, Parfrey LWL, Clemente JC, Gevers D and Knight R. 2012. “Experimental and Analytical Tools for Studying the Human Microbiome.” *Nature Reviews Genetics* 13 (1): 47–58.

Ladau L, Sharpton TJ, Finucane MM, Jospin G, Kembel SW, O’Dwyer J, Koeppel AF, Green JL and Pollard KS. 2013. “Global Marine Bacterial Diversity Peaks at High Latitudes in Winter.” *The ISME J*7 (9): 1669–1677.

Lebedeva EV, Hatzenpichler R, Pelletier E, Schuster N, Hauzmayer S, Aleksandr Bulaev A, Nadezhda V Grigor’eva, Galushko A, Schmid M, Palatinszky M, Le Paslier D, Daims H, Wagner M. 2013. “Enrichment and Genome Sequence of the Group I.1a Ammonia-Oxidizing Archaeon ‘Ca. *Nitrosotenuis Uzonensis*’ Representing a Clade Globally Distributed in Thermal Habitats.” *PloS One* 8 (11): e80835.

Lee CK, Barbier BA, Bottos EM, McDonald IR and Cary SC. 2012. “The Inter-Valley Soil Comparative Survey: The Ecology of Dry Valley Edaphic Microbial Communities.” *The ISME J.* 6 (5): 1046–1057.

Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol GW, Prosser JI, Schuster SC and Schleper C. 2006. “Archaea Predominate among Ammonia-Oxidizing Prokaryotes in Soils.” *Nature* 442: 806–809.

Leinonen R, Akhtar R, Birney E, Bower L, Cerdeno-Tárraga A, Cheng Y, Cleland I, Faruque N, Goodgame N, Gibson R, Hoad G, Jang M, Pakseresht N, Plaister S, Radhakrishnan R, Reddy K, Sobhany S, Hoopen PT, Vaughan R, Zalunin V and Cochrane G. 2011. “The European Nucleotide Archive.” *Nucleic Acids Research* 39 (Database issue): D28–D31.

Li W and Godzik A. 2006. “Cd-Hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences.” *Bioinformatics* 22 (13): 1658–1659.

Liles MR, Manske BF, Bintrim SB, Handelsman J and Goodman RM. 2003. “A Census of rRNA Genes and Linked Genomic Sequences within a Soil Metagenomic Library.” *Applied and Environmental Microbiology* 69 (5): 2684–2691.

Lipson DA and Schmidt SK. 2004. “Seasonal Changes in an Alpine Soil Bacterial Community in the Colorado Rocky Mountains.” *Applied and Environmental Microbiology* 70 (5): 2867–2879.

Liu H, Xu Y, Ma Y and Zhou P. 2000. “Characterization of Micrococcus Antarcticus Sp. Nov., a Psychrophilic Bacterium from Antarctica.” *International Journal of Systematic and Evolutionary Microbiology* 50 (2): 715–719.

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L and Law M. 2012. “Comparison of next-Generation Sequencing Systems.” *Journal of Biomedicine & Biotechnology* 2012: 1-11.

Liu WT and Marsh TL. 1997. “Characterization of Microbial Diversity by Determining Terminal Restriction Fragment Length Polymorphisms of Genes Encoding 16S rRNA.” *Applied and Environmental Microbiology* 63 (11): 4516–4522.

Lozupone CA, Hamady M, Kelley ST and Knight R. 2007. “Quantitative and Qualitative Beta Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities.” *Applied and Environmental Microbiology* 73 (5): 1576–1585.

Lozupone, CA and Knight R. 2005. “UniFrac: A New Phylogenetic Method for Comparing Microbial Communities.” *Applied and Environmental Microbiology* 71 (12): 8228–8235.

Magalhães, C., Machado A., Frank-Fahle, B., Lee, CK., and Cary, SC. 2014. “The Ecological Dichotomy of Ammonia-Oxidizing Archaea and Bacteria in the Hyper-Arid Soils of the Antarctic Dry Valleys.” *Frontiers in Microbiology* 5 (10): 1-11.

Magalhães, C., Stevens, MI., Stevens, I., Cary, SC., Ball, BA, Storey, VC, Wall, DH, Turk, R and Ruprecht, U. 2012. “At Limits of Life: Multidisciplinary Insights Reveal Environmental Constraints on Biotic Diversity in Continental Antarctica.” *PloS One* 7 (9): e44578.

Magalhães, C., Machado, A., and Bordalo A. 2009. “Temporal Variability in the Abundance of Ammonia-Oxidizing Bacteria vs. Archaea in Sandy Sediments of the Douro River Estuary, Portugal.” *Aquatic Microbial Ecology* 56 (July): 13–23.

Maidak, BL, Olsen, GJ and Larsen, N. 1996. “The Ribosomal Database Project (RDP).” *Nucleic Acids Research* 24 (1): 82–85.

Marchant, DR. and Head, JW. 2007. “Antarctic Dry Valleys: Microclimate Zonation, Variable Geomorphic Processes, and Implications for Assessing Climate Change on Mars.” *Icarus* 192 (1): 187–222.

Margesin, R., and Miteva, V. 2011. “Diversity and Ecology of Psychrophilic Microorganisms.” *Research in Microbiology* 162 (3). Elsevier Masson SAS: 346–361.

Marshall, WA, and Chalmers, MO. 1997. “Airborne Dispersal of Antarctic Terrestrial Algae and Cyanobacteria.” *Ecography* 20: 585–594.

Martens-Habbena, W., Berube, PM., Urakawa,H., de la Torre, JR., Stahl, DA.. 2009. “Ammonia Oxidation Kinetics Determine Niche Separation of Nitrifying Archaea and Bacteria.” *Nature* 461 (7266): 976–979.

McDonald, D., Price, MN, Goodrich, J., Nawrocki, EP., DeSantis, TZ, Probst, A., Andersen, GL, Knight, R., Hugenholtz, P. 2012. “An Improved Greengenes Taxonomy with Explicit Ranks for Ecological and Evolutionary Analyses of Bacteria and Archaea.” *The ISME Journal* 6 (3): 610–618.

McGowan, HA., Neil, DT and Speirs, JC. 2014. “A Reinterpretation of Geomorphological Evidence for Glacial Lake Victoria, McMurdo Dry Valleys, Antarctica.” *Geomorphology* 208 (3). Elsevier B.V. 200–206.

McTavish, HJAF, Fuchs, JA., and Hooper, AB. 1993. “Sequence of the Gene Coding for Ammonia Monooxygenase in *Nitrosomonas europaea*.” *Journal of Bacteriology* 175 (8). 2436–2444.

Method of the year 2013, *Nature Methods* 11,1 (2014)

Mocali, S., and Benedetti A. 2010. “Exploring Research Frontiers in Microbiology: The Challenge of Metagenomics in Soil Microbiology.” *Research in Microbiology* 161 (6). Elsevier Masson SAS: 497–505.

Monteiro, M., Séneca, J., and Magalhães, C. 2014. “The History of Aerobic Ammonia Oxidizers: From the First Discoveries to Today.” *Journal of Microbiology* 52 (7): 537–547.

Mori, H., Maruyama, F., Kato, H., Toyoda, A., Dozono, A., Ohtsubo, Y., Nagata, Y., Fujiyama, A., Tsuda, M., and Kurokawa, K. 2014. "Design and Experimental Application of a Novel Non-Degenerate Universal Primer Set That Amplifies Prokaryotic 16S rRNA Genes with a Low Possibility to Amplify Eukaryotic rRNA Genes." *DNA Research : An International Journal for Rapid Publication of Reports on Genes and Genomes* 21 (2): 217–227.

Morita, RY. 1975. "Psychrophilic Bacteria." *Bacteriological Reviews* 39 (2): 144–167.

Mulder, A, van de Graaf, AA., Robertson, LA., Kuenen, JG. 1995. "Annaerobic Ammonium Oxidation Discovered in a Denitrifying Fluidized Bed Reactor." *FEMS Microbiology Letters* 16: 177–184.

Muyzer, G, De Wall, EC., and Uitterlinden, AG. 1993. "Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA." *Applied and Environmental Microbiology* 59 (3): 695–700.

Navas-Molina, JA, Peralta-Sánchez, JM., González, A., McMurdie, PJ, Vázquez-Baeza, Y., Xu, Z., Ursel, LK et al. 2013. *Advancing Our Understanding of the Human Microbiome Using QIIME. Methods in Enzymology*. 1st ed. Vol. 531. Elsevier Inc.

Nicol, GW, and Schleper C. 2006. "Ammonia-Oxidising Crenarchaeota: Important Players in the Nitrogen Cycle?" *Trends in Microbiology* 14 (5): 207–212.

Nicol, GW, Tscherko, D., Embley, TM., Prosser, JI. 2005. "Primary Succession of Soil Crenarchaeota across a Receding Glacier Foreland." *Environmental Microbiology* 7 (3): 337–347.

Niederberger, TD, McDonald, IR, Hacker, AL, Soo, RM., Barrett, JE., Wall, DH and Cary, SC.. 2008. "Microbial Community Composition in Soils of Northern Victoria Land, Antarctica." *Environmental Microbiology* 10 (7): 1713–1724.

Nikolaki, S., and Tsiamis, G.. 2013. "Microbial Diversity in the Era of Omic Technologies." *BioMed Research International* 2013 (1): 958719.

Nirenberg, M., Caskey, T., Marshall, R., Brimacombe, R., Kellogg, D., Doctor, B., Hatfield, D. et al 1966. "The RNA Code and Protein Synthesis." *Cold Spring Harbor Symposia on Quantitative Biology* 31 (January): 11–24.

Norton, JM, Alzerreca, JJ, Suwa, Y., and Klotz, MG. 2002. "Diversity of Ammonia Monooxygenase Operon in Autotrophic Ammonia-Oxidizing Bacteria." *Archives of Microbiology* 177 (2): 139–149.

Nordberg H., et al. 2014 The genome portal of the Department of Energy Joint Genome Institute: 2014 updates.*Nucleic Acids Research* ,42(1), pp. 26-31

Nugroho, RA, Röling, WFM, Laverman, AM, Zoomer, HR and Verhoef, HA. 2005. "Presence of Nitrosospira Cluster 2 Bacteria Corresponds to N Transformation Rates in Nine Acid Scots Pine Forest Soils." *FEMS Microbiology Ecology* 53 (3): 473–481.

Nunan, N, Wu, K., Young, IM, Crawford, JW, and Ritz, K. 2002. "In Situ Spatial Patterns of Soil Bacterial Populations, Mapped at Multiple Scales, in an Arable Soil." *Microbial Ecology* 44 (4): 296–305.

Nylen, TH. 2004. "Climatology of Katabatic Winds in the McMurdo Dry Valleys, Southern Victoria Land, Antarctica." *Journal of Geophysical Research* 109 (D3): D03114.

Ochsenreiter, T., Pfeifer, F. and Schleper, C. 2002. "Diversity of Archaea in Hypersaline Environments Characterized by Molecular-Phylogenetic and Cultivation Studies." *Extremophiles* 6: 267–274.

Ochsenreiter, T., Selezi, D., Quaiser,A., Bonch-Osmolovskaya, L and Schleper, C. 2003. “Diversity and Abundance of Crenarchaeota in Terrestrial Habitats Studied by 16S RNA Surveys and Real Time PCR.” *Environmental Microbiology* 5 (9). Blackwell Science Ltd: 787–797.

Okano, Y., Hristova, KR., Leutenegger, CM, Jackson, LE., Denison, RF., Gebreyesus, B., Lebauer, D. and Scow, KM. 2004. “Application of Real-Time PCR to Study Effects of Ammonium on Population Size of Ammonia-Oxidizing Bacteria in Soil.” *Applied and Environmental Microbiology* 70 (2): 1008–1016.

Olsen, GJ, Lane, DJ, Giovannoni, SJ, Pace, NR, and Stahl, DA. 1986. “Microbial Ecology and Evolution: A Ribosomal RNA Approach.” *Annual Review of Microbiology* 40 (1): 337–365.

Ondov, BD. Bergman, NH and Phillippy, AM. 2011. “Interactive Metagenomic Visualization in a Web Browser.” *BMC Bioinformatics* 12 (1). BioMed Central Ltd: 385.

Pace, NR. 1997. “A Molecular View of Microbial Diversity and the Biosphere.” *Science* 276 (5313). American Association for the Advancement of Science: 734–740.

Pace, NR, Sapp, J. and Goldenfeld, N. 2012. “Phylogeny and beyond: Scientific, Historical, and Conceptual Significance of the First Tree of Life.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (4): 1011–1018.

Parameswaran, P., Jalili, R., Tao, L., Shokralla, S, Gharizadeh, B., Ronaghi M. and Fire, AZ. “A Pyrosequencing-Tailored Nucleotide Barcode Design Unveils Opportunities for Large-Scale Sample Multiplexing.” *Nucleic Acids Research* 35 (19): e130.

Pesaro, M., Widmer, F., Nicollier, G., and Zefer, J. 2003. “Effects of Freeze–thaw Stress during Soil Storage on Microbial Communities and Methidathion Degradation.” *Soil Biology and Biochemistry* 35 (8): 1049–1061.

Pester, M., Rattei, T., Flechl, S., Gröngröft, A., Richter, A., Overmann, J., Reinhold-hurek, B., Loy, Al. and Wagner, M. 2012. “amoA-Based Consensus Phylogeny of Ammonia-Oxidizing Archaea and Deep Sequencing of amoA Genes from Soils of Four Different Geographic Regions.” *Environmental Microbiology* 14 (2): 525–539.

Pointing, SB, Chan, Y., Lacap, DC, Lau, MCY , Jurgens, JA and Farrell, RL 2009. “Highly Specialized Microbial Diversity in Hyper-Arid Polar Desert.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (47): 19964–19969.

Poplawski, AB, Mårtensson, L., Wartiainen,I., and Rasmussen U.. 2007. “Archaeal Diversity and Community Structure in a Swedish Barley Field: Specificity of the EK510R/(EURY498) 16S rDNA Primer.” *Journal of Microbiological Methods* 69 (1): 161–173.

Poretsky, R, Rodriguez-R, LM, Luo, C.,Tsementzi, D., and Konstantinidis, KT. 2014. “Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics.” *PloS One* 9 (4): e93827.

Preston, CM, Ying Wu, KE, Molinskit, TF and Delong, EF. 1996. “A Psychrophilic Crenarchaeon Inhabits a Marine Sponge :” *Proceedings of the National Academy of Sciences* 93 (6): 6241–6246.

Price, MN, Dehal, PS and Arkin, AP. 2010. “FastTree 2--Approximately Maximum-Likelihood Trees for Large Alignments.” *PloS One* 5 (3): e9490.

Pruesse, E., Peplies, J., and Glöckner, FO. 2012. “SINA: Accurate High-Throughput Multiple Sequence Alignment of Ribosomal RNA Genes.” *Bioinformatics (Oxford, England)* 28 (14): 1823–1829.

Purkhold, U, Pommerening-Röser, A., Juretschko, S., Schmid, MC, Koops, HP and Wagner, M. 2000. "Phylogeny of All Recognized Species of Ammonia Oxidizers Based on Comparative 16S rRNA and amoA Sequence Analysis: Implications for Molecular Diversity Surveys." *Applied and Environmental Microbiology* 66 (12): 5368–5382.

Purkhold, U, Wagner, M, Timmermann, G., Pommerening-Röser, A. and. Koops, HP. 2003. "16S rRNA and amoA-Based Phylogeny of 12 Novel Betaproteobacterial Ammonia-Oxidizing Isolates: Extension of the Dataset and Proposal of a New Lineage within the Nitrosomonads." *International Journal of Systematic and Evolutionary Microbiology* 53 (5): 1485–1494.

Quast, C., Pruesse, E. , Yilmaz, EP., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckne, FO. 2013. "The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic Acids Research* 41 (Database issue): D590–596.

Radax, R., Hoffmann, F., Rap, HT, Leininger, S. and Schleper, C. 2012. "Ammonia-Oxidizing Archaea as Main Drivers of Nitrification in Cold-Water Sponges." *Environmental Microbiology* 14 (4): 909–923.

Rampelotto, PH. 2010. "Resistance of Microorganisms to Extreme Environmental Conditions and Its Contribution to Astrobiology." *Sustainability* 2 (6): 1602–1623.

Reeder, J and Knight, R. 2010. "Rapidly Denoising Pyrosequencing Amplicon Reads by Exploiting Rank-Abundance Distributions." *Nature Methods* 7 (9): 2009–2011.

Richter, I. Herbold, CW., Lee, CK., McDonald, IR, Barret, JE, Cary, SC. 2014. "Influence of Soil Properties on Archaeal Diversity and Distribution in the McMurdo Dry Valleys, Antarctica." *FEMS Microbiology Ecology* 89 (2): 347–359.

Riesenfeld, CS, Schloss, PD and Handelsman, J. 2004. "Metagenomics: Genomic Analysis of Microbial Communities." *Annual Review of Genetics* 38 (January): 525–552.

Rinke, C., Schwientek, P., Sczyrba, A. Ivanova, NN, Anderson, IJ, Cheng, JF, Darling, A. et al 2013. "Insights into the Phylogeny and Coding Potential of Microbial Dark Matter." *Nature* 499 (7459). Nature Publishing Group: 431–437.

Rocha, EP. 2008. "Evolutionary Patterns in Prokaryotic Genomes." *Current Opinion in Microbiology* 11 (5): 454–460.

Rothschild, LJ, and Mancinelli RL 2001. "Life in Extreme Environments." *Nature* 409 (6823): 1092–1101.

Rotthauwe, JH, de Boer, W and Liesack, W. 1995. "Comparative Analysis of Gene Sequences Encoding Ammonia Monooxygenase of *Nitrosospira Sp. AHB1* and *Nitrosolobus Multiformis C-71*." *FEMS Microbiology Letters* 133 (1-2): 131–135.

Rotthauwe, JH, Witzel, KP and Liesack, W. 1997. "The Ammonia Monooxygenase Structural Gene *amoA* as a Functional Marker: Molecular Fine-Scale Analysis of Natural Ammonia-Oxidizing Populations." *Applied and Environmental Microbiology* 63 (12): 4704–4712.

Sanger, F, and Coulson, AR 1975. "A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase." *Journal of Molecular Biology* 94 (3): 441–448.

Sanger, F, Nicklen, S and Coulson, R. 1977. "DNA Sequencing with Chain-Terminating." *Proceedings of the National Academy of Sciences* 74 (12): 5463–5467.

Santoro, AE, Francis C., de Sieyes, NR and Boehm, AB 2008. “Shifts in the Relative Abundance of Ammonia-Oxidizing Bacteria and Archaea across Physicochemical Gradients in a Subterranean Estuary.” *Environmental Microbiology* 10 (4): 1068–1079.

Satyanarayana, T., Raghukumar, C. and Shivaji, S. 2005. “Extremophilic Microbes: Diversity and Perspectives.” *Current Science* 89 (1):78-90

Schafer, G. 2004. “Extremophilic Archaea and Bacteria.” *Journal of Bioenergetics and Biomembranes* 36 (1): 3–4.

Scheu, S., Ruess, L. & Bonkowski, M., 2005. Interactions Between Microorganisms and Soil Micro- and Mesofauna, in Microorganisms in *Soils: Roles in Genesis and Functions, part IV*, (eds: A. Varma & F. Buscot), 253-275

Schleper, C., Nicol, G. W., 2010, Ammonia-Oxidising Archaea – Physiology, Ecology and Evolution. In: *Advances in Microbial Physiology*, Vol 57, (eds: R. K. Poole), Oxford: Academic Press; 2010, pp. 1-41

Schloss, PD, and Handelsman, J. 2004. “Status of the Microbial Census.” *Microbiology and Molecular Biology Reviews : MMBR* 68 (4): 686–691.

Schloss, PD, and Westcott, SL 2011. “Assessing and Improving Methods Used in Operational Taxonomic Unit-Based Approaches for 16S rRNA Gene Sequence Analysis.” *Applied and Environmental Microbiology* 77 (10): 3219–3226.

Schloss, PD, Westcott, SL, Ryabin, T., Hall, JR, Hartmann, M., Hollister, EB, Lesniewski, RA et al 2009 “Introducing Mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities.” *Applied and Environmental Microbiology* 75 (23): 7537–7541.

Segata, N. Boernigen, S., Tickle, TL, Morgan, XC, Garret, WS, Huttenhower, C. 2013. “Computational Meta’omics for Microbial Community Studies.” *Molecular Systems Biology* 9 (666). Nature Publishing Group: 666.

Seybold, CA, Harms, DS, Balks, M, Aislabie, J., Paetzold, RF, Kimble, J. and Sletten, R. 2009. “Soil Climate Monitoring Project in the Ross Island Region of Antarctica.” *Soil Survey Horizons*, 52–57.

Shtarkman, YM, Koçer, ZK, Edgar, R., Veerapaneni, RS, D’Elia, T., Morris, PF., Rogers, SO. “Subglacial Lake Vostok (Antarctica) Accretion Ice Contains a Diverse Set of Sequences from Aquatic, Marine and Sediment-Inhabiting Bacteria and Eukarya.” *PloS One* 8 (7): e67221.

Smalla, K., Oros-Sichler, M., Milling, A., Heuer, H., Baumgarte, S., Becker, R., Neuber, G., Kropf, S., Ulrich, A., Tebbe, CC. 2007. “Bacterial Diversity of Soils Assessed by DGGE, T-RFLP and SSCP Fingerprints of PCR-Amplified 16S rRNA Gene Fragments: Do the Different Methods Provide Similar Results?” *Journal of Microbiological Methods* 69 (3): 470–479.

Smith, JJ, Tow, LA, Stafford, W., Cary, SC and Cowan, DA. 2006. “Bacterial Diversity in Three Different Antarctic Cold Desert Mineral Soils.” *Microbial Ecology* 51 (4): 413–421.

Sogin, ML, Morrison, HG, Huber, JA., Welch, DM., Huse, SM, Neal, PR, Arrieta, JM , and Herndl, GJ 2006. “Microbial Diversity in the Deep Sea and the Underexplored ‘Rare Biosphere’.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (32): 12115–12120.

Sokol, ER, Herbold, CW, Lee,CK, Cary, SC and Barrett, JE. 2013. “Local and Regional Influences over Soil Microbial Metacommunities in the Transantarctic Mountains.” *Ecosphere* 4 (November): 1–24.

Song, Y., Zhang, X., Ma, B., Chang, S., Gong, J. 2014. “Biochar Addition Affected the Dynamics of Ammonia Oxidizers and Nitrification in Microcosms of a Coastal Alkaline Soil.” *Biology and Fertility of Soils* 50 (2). Springer Berlin Heidelberg: 321–332.

Spang, A, Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., Stahl, DAl Wagner, M., and Schleper, C. 2010. “Distinct Gene Set in Two Different Lineages of Ammonia-Oxidizing Archaea Supports the Phylum Thaumarchaeota.” *Trends in Microbiology* 18 (8). Elsevier: 331–340.

Spang, A, Poehlein, A., Offre, P. Zumbrägel, S., Haider, S., Rychlik, N. Boris Nowka, et al. 2012. “The Genome of the Ammonia-Oxidizing Candidatus Nitrososphaera Gargensis: Insights into Metabolic Versatility and Environmental Adaptations.” *Environmental Microbiology* 14 (12): 3122–3145.

Stackebrandt, E. & Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiology Today* 33, 152–155.

Stackebrandt, E., and Goebel, BM 1994. “Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology.” *International Journal of Systematic Bacteriology* 44 (4): 846–849.

Stackebrandt, E. 2011. “Towards a Strategy to Enhance Access to Microbial Diversity.” *International Journal of Systematic and Evolutionary Microbiology* 61 (Pt 3): 479–481.

Stahl, DA, and de la Torre JR. 2012. “Physiology and Diversity of Ammonia-Oxidizing Archaea.” *Annual Review of Microbiology* 66. Department of Civil and Environmental Engineering, University of Washington, Seattle, 98195-2700, USA. dastahl@u.washington.edu: 83–101.

Stein, LD. 2010. “The Case for Cloud Computing in Genome Informatics.” *Genome Biology* 11 (5): 207.

Stephen, J.R. et al. 1996. Molecular diversity of soil and marine 16S rRNA gene sequences related to beta-subgroup Molecular Diversity of Soil and Marine 16S rRNA Gene Sequences Related to ␣ -Subgroup Ammonia-Oxidizing Bacteria. *Applied and Environmental Microbiology*, 62(11), pp. 4147-4154

Sun, HY, Deng, SP and Raun, WR. 2004. “Bacterial Community Structure and Diversity in a Century-Old Manure-Treated Agroecosystem.” *Applied and Environmental Microbiology* 70 (10): 5868–5874.

Talarico, F.M. & Kleinschmidt, G., 2008. The Antarctic Continent in Gondwanaland: A Tectonic Reviewand Potential Research Targets for Future Investigations, Antarctic Climate Evolution, 8, pp.257–308

Tamura, K, Stecher,G. Peterson, D., Filipski, A. and Kumar, S. 2013. “MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0.” *Molecular Biology and Evolution* 30 (12): 2725–2729.

Tateno, Y, and Fukami-Kobayashi, K. 1998. “DNA Data Bank of Japan at Work on Genome Sequence Data.” *Nucleic Acids Research* 26 (1): 16–20.

Thamdrup, B.. 2012. “New Pathways and Processes in the Global Nitrogen Cycle.” *Annual Review of Ecology, Evolution, and Systematics* 43 (1): 407–428.

Thomas, F., Hehemann, JH, Etienne R., Czjzek, M. and Michel, G. 2011. “Environmental and Gut Bacteroidetes: The Food Connection.” *Frontiers in Microbiology* 2 (5): 1-16.

Thompson, JD. Higgins, DG and. Gibson, TJ. 1994. "CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice." *Nucleic Acids Research* 22 (22): 4673–4680.

Tindall, B J, Rosselló-Móra, R., Busse, HJ, Ludwig, W, and Kämpfer, P.. 2010. "Notes on the Characterization of Prokaryote Strains for Taxonomic Purposes." *International Journal of Systematic and Evolutionary Microbiology* 60 (Pt 1): 249–66.

Torsvik, VJ G, and Daae. 1990. "High Diversity in DNA of Soil Bacteria." *Applied and Environmental Microbiology* 56 (3): 782–787.

Torsvik, V, and Øvreås, L.. 2002. "Microbial Diversity and Function in Soil: From Genes to Ecosystems." *Current Opinion in Microbiology* 5 (3): 240–245.

Tourna, M. and Stieglmeier, M.. 2011. "*Nitrososphaera Viennensis*, an Ammonia Oxidizing Archaeon from Soil." *Proceedings of the National Academy of Sciences* 1–6.

Treusch, AH, Leininger, S. Kletzin, A., Schuster, SC, Klenk, HP, and Schleper, C.. 2005. "Novel Genes for Nitrite Reductase and Amo-Related Proteins Indicate a Role of Uncultivated Mesophilic Crenarchaeota in Nitrogen Cycling." *Environmental Microbiology* 7 (12): 1985–1995.

Van Elsas, J.D., and Boersma FGH.. 2011. "A Review of Molecular Methods to Study the Microbiota of Soil and the Mycosphere." *European Journal of Soil Biology* 47 (2). Elsevier Masson SAS: 77–87.

Vázquez-Baeza, Y,Pirrung, M., Gonzalez, A. and Knight, R. 2013. "EMPeror: A Tool for Visualizing High-Throughput Microbial Community Data." *GigaScience* 2 (1): 16.

Venter, JC, Remington, K. and Heidelberg, JF. 2004. "Environmental Genome Shotgun Sequencing of the Sargasso Sea." *Science* 304 (April): 66–75.

Ventura, M, Canchaya, C,Tauch, A, Chandra, G. Fitzgerald, GF, Chater, KF and Sinderen, D. 2007. "Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum." *Microbiology and Molecular Biology Reviews : MMBR* 71 (3): 495–548.

Vincent, F. 2004. "Evolutionary Origins of Antarctic Microbiota: Invasion, Selection and Endemism." *Antarctic Science* 12 (03): 374–385.

Vos, M, Quince, C., Pijl, AS, Hollander, M., Kowalchuk, GA. 2012. "A Comparison of rpoB and 16S rRNA as Markers in Pyrosequencing Studies of Bacterial Diversity." *PloS One* 7 (2): e30600.

Voytek, M., Priscu, J. and Ward, B., 1999. The distribution and relative abundance of ammonia-oxidizing bacteria in lakes of the McMurdo Dry Valley, Antarctica, *Hydrobiologia*, 401, pp. 113-130

Voytek, M. and Ward, B., 1995. Detection of ammonium-oxidizing bacteria of the beta-subclass of the class Proteobacteria in aquatic samples with the PCR. *Applied and environmental microbiology*, 61(4)., pp. 1444-1450

Wagner, S. C. (2011) Biological Nitrogen Fixation. *Nature Education Knowledge* 3(10):15

Walker, CB, de la Torre, JR, Klotz, V, Urakawa, H, Pinel, N., Arp, DJ, Brochier-Armanet C., et al 2010. "*Nitrosopumilus Maritimus* Genome Reveals Unique Mechanisms for Nitrification and Autotrophy in Globally Distributed Marine Crenarchaea." *Proceedings of the National Academy of Sciences of the United States of America* 107 (19): 8818–8823.

Wall, DH. 2005. “Biodiversity and Ecosystem Functioning in Terrestrial Habitats of Antarctica.” *Antarctic Science* 17 (04): 523.

Wall, DH., Adams, BJ, Barret, JE, Hopking, DW, and Virginia RA. 2006. “A Synthesis of Soil Biodiversity and Ecosystem Functioning in Victoria Land, Antarctica.” *Soil Biology and Biochemistry* 38 (10): 3001–3002.

Wang, Y., Qian, PY. 2009. “Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies.” *PloS One* 4 (10): e7401.

Ward, D., Weller, R. and Bateson, M., 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community, 75, pp.105-116

Waterhouse, AM, Procter, JB, Martin, DMMA, Clamp, M., Barton, GF 2009. “Jalview Version 2--a Multiple Sequence Alignment Editor and Analysis Workbench.” *Bioinformatics (Oxford, England)* 25 (9): 1189–1191.

Watson, JD, and Crick, FHC. 1953. “Molecular Structure of Nucleic Acids.” *Nature* 171: 737–738.

Webster-Brown, J., Gall, J., Gibson, J., Wood, S., and Hawes, I. 2010. “The Biogeochemistry of Meltwater Habitats in the Darwin Glacier Region (80°S), Victoria Land, Antarctica.” *Antarctic Science* 22 (06): 646–661.

Weisburg, WG, Barns, SM, Pelletier, DA and Lane, DJ, 1991. “16S Ribosomal DNA Amplification for Phylogenetic Study.” *Journal of Bacteriology* 173 (2): 697–703.

Wilson, KH, Blitchington, RB, and Greene, RC. 1990. “Amplification of Bacterial 16S Ribosomal DNA with Polymerase Chain Reaction.” *Journal of Clinical Microbiology* 28 (9).

Winogradsky, S. (1890): Sur les organismes de la nitrification. In: *Comptes rendus de séances l’Academie des Sciences*. Bd., 110, 1013–1016.

Woese, C R, and Fox, G.E.. 1977. “Phylogenetic Structure of the Prokaryotic Domain: The Primary Kingdoms.” *Proceedings of the National Academy of Sciences of the United States of America* 74 (11): 5088–5090.

Woese, C. R, Kandlert, O., Wheelis, M. L. 1990. “Towards a Natural System of Organisms : Proposal for the Domains Archaea, Bacteria and Eucarya.” *Proceedings of the National Academy of Sciences* 87 (6): 4576–4579.

Wood, S. A., Mountfort, D., Selwood, A., I., Holland, P. T., Puddick, J and Cary, S.C. “Widespread Distribution and Identification of Eight Novel Microcystins in Antarctic Cyanobacterial Mats.” *Applied and Environmental Microbiology* 74 (23): 7243–7251.

Yergeau, E., Kang, S., He, Z., Zhou, J., Kowalchuk, G.A. 2007a. “Functional Microarray Analysis of Nitrogen and Carbon Cycling Genes across an Antarctic Latitudinal Transect.” *The ISME Journal* 1 (2): 163–179.

Yergeau, E., Newsham, K. K. Pearce, D. A., Kowalchuk, G.A. 2007b. “Patterns of Bacterial Diversity across a Range of Antarctic Terrestrial Habitats.” *Environmental Microbiology* 9 (11): 2670–2682.

Yung, C., Chan, C.M.C., Lacap, D.C, Pérez-Ortega, A.L, Lee, C. K., Cary, S.C. and Pointing, S. B. 2014. “Characterization of Chasmoendolithic Community in Miers Valley, McMurdo Dry Valleys, Antarctica.” *Microbial Ecology* 68 (2): 351–359.

Zeglin, L. H., Sinsabaugh, R. L., Barret, J. E., Gooseff, M. N., Michael N., and Takacs-Vesbach, C. D.. 2009. “Landscape Distribution of Microbial Activity in the McMurdo Dry Valleys: Linked Biotic Processes, Hydrology, and Geochemistry in a Cold Desert Ecosystem.” *Ecosystems* 12 (4): 562–573.

Zhalnina, K., de Quadros, P.D., Camargo, F. A., Troplett, E.W.. “Drivers of Archaeal Ammonia-Oxidizing Communities in Soil.” *Frontiers in Microbiology* 3 (June): 210:1-9

Zuckerkandl, E, and L Pauling. 1965. “Molecules as Documents of Evolutionary History.” *Journal of Theoretical Biology* 8 (2): 357–366.

Supplementary information

Appendix 1. General QIIME workflow

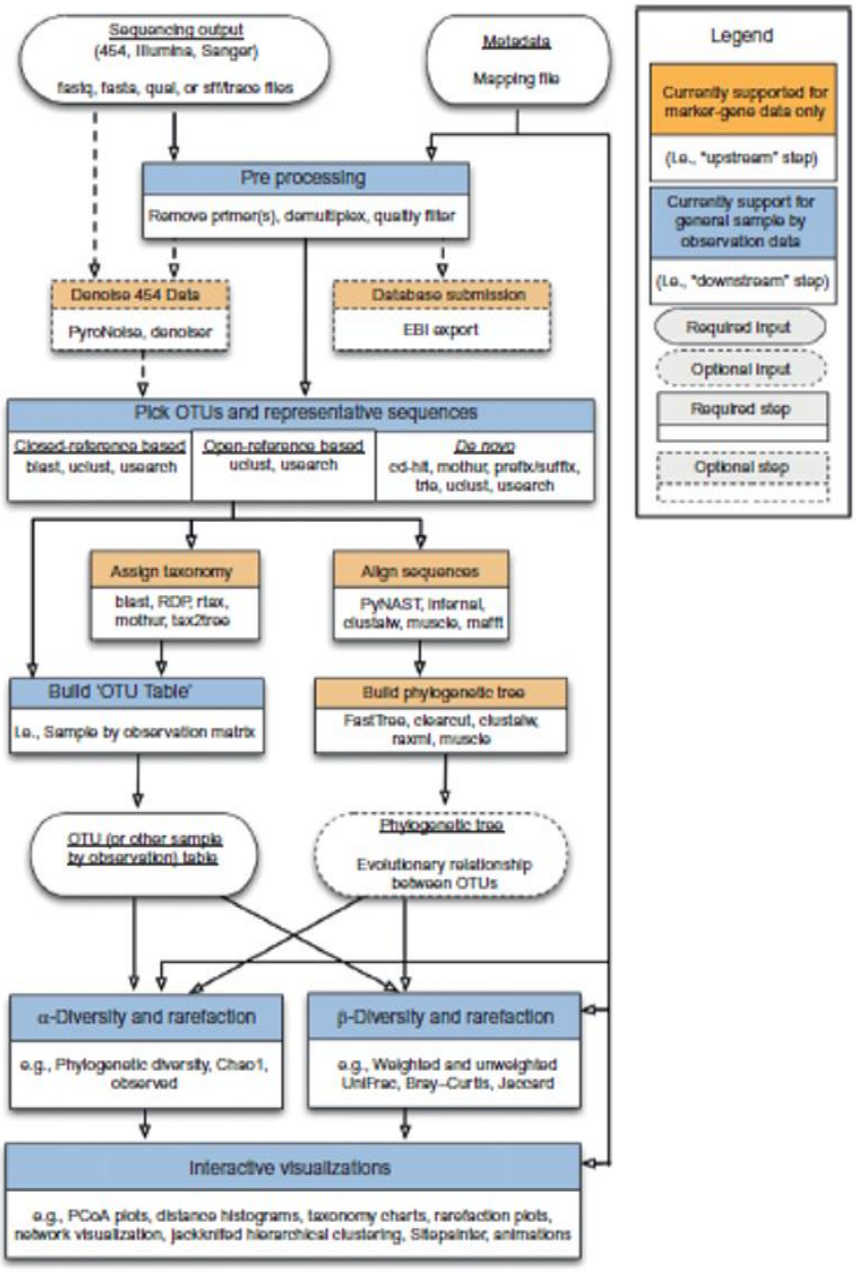


Figure 1. General workflow for NGS data analysis based on QIIME. Adopted from www.wernerlab.org/teaching/qiime/overview/g

Appendix 2. Class- based taxonomy on QIIME

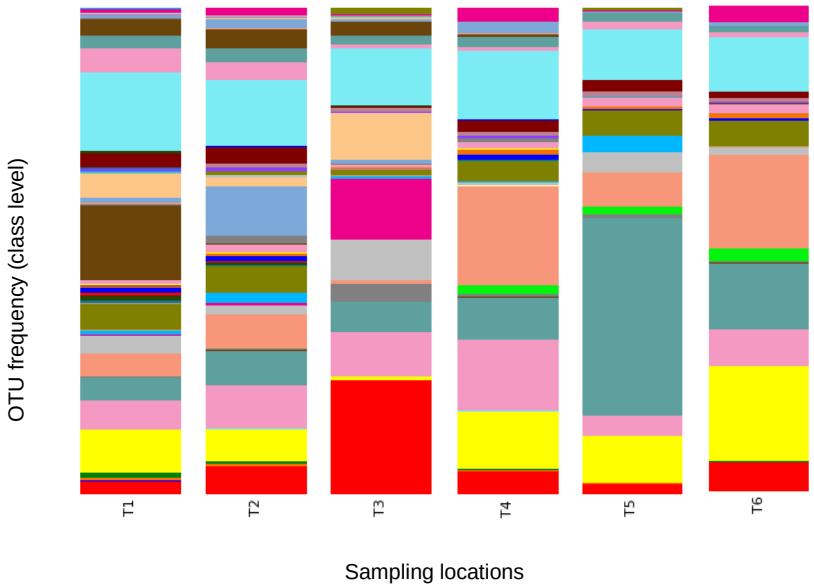


Figure 1. Relative abundance of class-affiliated OTUs per sampling point. Color code is presented in Table 1.

Table 1. Frequency of OTU abundance at the class-level. All values are percentages.

	Classes	Total	T1	T2	T3	T4	T5	T6
	Unassigned;Other;Other	7.40	2.71	5.77	23.34	4.76	2.19	5.91
	k__Bacteria;p__Acidobacteria;c__AT-s54	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6	0.30	0.52	0.59	0.19	0.18	0.11	0.19
	k__Bacteria;p__Acidobacteria;c__Solibacteres	0.40	1.26	0.47	0.00	0.30	0.01	0.16
	k__Bacteria;p__Acidobacteria;c__Sva0725	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria]	9.40	8.67	6.36	0.73	11.76	9.75	19.43
	k__Bacteria;p__Acidobacteria;c__iii1-8	0.10	0.12	0.33	0.03	0.09	0.00	0.03
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia	8.40	5.82	8.89	9.08	14.75	4.19	7.67
	k__Bacteria;p__Actinobacteria;c__Actinobacteria	13.40	4.81	6.96	6.26	8.50	40.57	13.52
	k__Bacteria;p__Actinobacteria;c__MB-A2-108	0.10	0.10	0.34	0.00	0.23	0.01	0.17
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria	0.90	0.03	0.16	3.60	0.49	0.62	0.36
	k__Bacteria;p__Actinobacteria;c__Rubrobacteria	1.10	0.10	0.18	0.11	1.87	1.68	2.55
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia	9.80	4.62	6.86	0.67	20.30	7.06	19.20
	k__Bacteria;p__Armatimonadetes;c__0319-6E2	0.00	0.01	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Armatimonadetes;c__Armatimonadia	0.00	0.00	0.01	0.00	0.03	0.01	0.02
	k__Bacteria;p__Armatimonadetes;c__[Fimbriimonadia]	0.00	0.11	0.07	0.00	0.01	0.00	0.00
	k__Bacteria;p__BRC1;c__PRR-11	0.00	0.01	0.00	0.03	0.08	0.07	0.01
	k__Bacteria;p__Bacteroidetes;c__At12OctB3	0.00	0.00	0.01	0.00	0.11	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia	3.30	3.57	1.86	8.22	0.52	3.97	1.53
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia	2.20	0.17	0.47	12.70	0.04	0.02	0.03
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia	1.20	0.94	2.03	0.29	0.29	3.33	0.16
	k__Bacteria;p__Bacteroidetes;c__[Rhodothermi]	0.10	0.13	0.07	0.43	0.01	0.01	0.02
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae]	4.40	5.37	5.46	1.10	4.12	5.38	5.24
	k__Bacteria;p__Chlamydiae;c__Chlamydiia	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chlorobi;c__OPB56	0.10	0.20	0.09	0.03	0.19	0.00	0.02
	k__Bacteria;p__Chlorobi;c__SJA-28	0.10	0.45	0.25	0.00	0.01	0.00	0.00

	Classes	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Chloroflexi;c__	0.00	0.01	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae	0.30	1.00	0.47	0.04	0.09	0.00	0.07
	k__Bacteria;p__Chloroflexi;c__C0119	0.00	0.09	0.03	0.00	0.10	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi	0.10	0.43	0.08	0.00	0.03	0.01	0.01
	k__Bacteria;p__Chloroflexi;c__Ellin6529	0.70	1.00	1.05	0.15	0.82	0.26	0.62
	k__Bacteria;p__Chloroflexi;c__Gitt-GS-136	0.60	0.30	0.68	0.13	1.09	0.50	0.83
	k__Bacteria;p__Chloroflexi;c__P2-11E	0.00	0.04	0.03	0.00	0.03	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__S085	0.10	0.38	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__TK10	0.10	0.06	0.05	0.00	0.18	0.03	0.22
	k__Bacteria;p__Chloroflexi;c__TK17	0.00	0.09	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia	1.30	0.73	1.77	0.47	1.33	1.86	1.81
	k__Bacteria;p__Chloroflexi;c__[Thermobacula]	0.00	0.00	0.00	0.00	0.02	0.02	0.03
	k__Bacteria;p__Cyanobacteria;c__	2.60	15.50	0.09	0.00	0.09	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Chloroplast	0.50	0.27	1.50	0.38	0.64	0.09	0.07
	k__Bacteria;p__Cyanobacteria;c__ML635J-21	0.00	0.04	0.03	0.00	0.03	0.08	0.02
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae	0.00	0.04	0.00	0.00	0.03	0.02	0.03
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae	2.00	1.20	10.19	0.83	0.01	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae	2.70	5.00	1.79	9.51	0.03	0.02	0.07
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia	0.10	0.20	0.09	0.03	0.04	0.00	0.02
	k__Bacteria;p__FBP;c__	0.00	0.01	0.05	0.00	0.00	0.06	0.02
	k__Bacteria;p__Fibrobacteres;c__Fibrobacteria	0.00	0.02	0.01	0.01	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli	0.10	0.01	0.36	0.03	0.00	0.00	0.01
	k__Bacteria;p__Firmicutes;c__Clostridia	0.00	0.03	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__BB34	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__GKS2-174	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__	0.10	0.05	0.82	0.00	0.02	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-1	0.40	0.58	0.94	0.13	0.37	0.01	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-3	0.70	0.00	0.53	0.79	0.91	0.84	0.83
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-5	0.10	0.11	0.17	0.05	0.08	0.01	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes	2.10	2.95	3.32	0.41	2.36	2.39	1.15

	Classes	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Nitrospirae;c__Nitrospira	0.10	0.60	0.13	0.02	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__Mb-NB09	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__OD1;c__SM2F11	0.00	0.00	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__OP11;c__WCHB1-64	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Phycisphaerae	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Planctomycetia	0.00	0.03	0.00	0.00	0.02	0.00	0.04
	k__Bacteria;p__Planctomycetes;c__vadinHA49	0.00	0.02	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria	12.90	15.98	13.75	11.91	14.04	10.37	11.31
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria	2.10	5.12	3.61	0.66	0.98	1.66	0.84
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria	1.90	2.38	2.78	1.74	1.90	1.62	1.27
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria	1.80	3.53	3.78	2.87	0.54	0.07	0.13
	k__Bacteria;p__Proteobacteria;c__TA18	0.10	0.33	0.21	0.00	0.17	0.00	0.04
	k__Bacteria;p__TM6;c__SJA-4	0.00	0.00	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__TM7;c__SC3	0.00	0.00	0.12	0.03	0.05	0.00	0.02
	k__Bacteria;p__TM7;c__TM7-1	1.10	0.30	1.98	0.71	2.44	0.29	0.61
	k__Bacteria;p__TM7;c__TM7-3	0.10	0.00	0.08	0.28	0.03	0.10	0.03
	k__Bacteria;p__Verrucomicrobia;c__Opitutae	0.00	0.05	0.07	0.02	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae	0.10	0.27	0.25	0.34	0.01	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Methylacidiphilae]	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae]	0.10	0.12	0.25	0.05	0.11	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria]	1.50	1.01	1.50	0.17	2.66	0.13	3.32
	k__Bacteria;p__WPS-2;c__	0.00	0.05	0.00	0.00	0.01	0.00	0.01
	k__Bacteria;p__WS3;c__PRR-12	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci	0.40	0.05	0.10	1.38	0.09	0.55	0.12

Appendix 3. Order-based taxonomy on QIIME

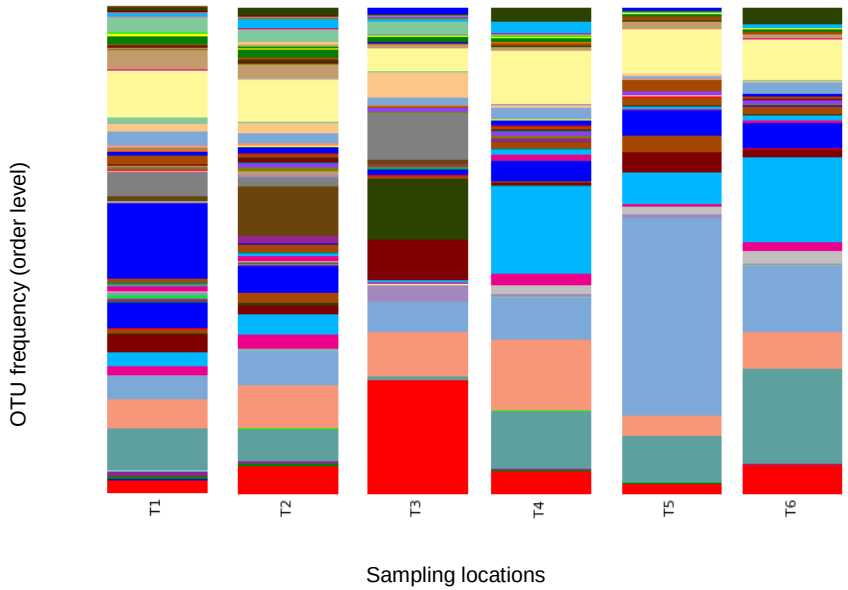


Figure 1. Relative abundance of order-affiliated OTUs per sampling point. Color code is presented in Table 1.

Table 1. Frequency of OTU abundance at the order-level. All values are percentages.

	Orders	Total	T1	T2	T3	T4	T5	T6
	Unassigned;Other;Other;Other	7.40	2.71	5.77	23.34	4.76	2.19	5.91
	k__Bacteria;p__Acidobacteria;c__AT-s54;o__	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__CCU21	0.00	0.06	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__iii1-15	0.30	0.45	0.57	0.19	0.18	0.11	0.19
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales	0.40	1.26	0.47	0.00	0.30	0.01	0.16
	k__Bacteria;p__Acidobacteria;c__Sva0725;o__Sva0725	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__DS-100	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__PK29	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__RB41	9.40	8.64	6.36	0.73	11.76	9.75	19.43
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__32-20	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__DS-18	0.10	0.11	0.30	0.03	0.09	0.00	0.03
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales	8.40	5.82	8.89	9.08	14.75	4.19	7.67
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales	13.40	4.81	6.96	6.26	8.50	40.57	13.52
	k__Bacteria;p__Actinobacteria;c__MB-A2-108;o__	0.00	0.00	0.00	0.00	0.00	0.00	0.02
	k__Bacteria;p__Actinobacteria;c__MB-A2-108;o__0319-7L14	0.10	0.10	0.34	0.00	0.23	0.01	0.15
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria;o__Euzebyales	0.80	0.03	0.16	3.32	0.49	0.62	0.36
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria;o__Nitriliruptorales	0.00	0.00	0.00	0.28	0.00	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Rubrobacteria;o__Rubrobacterales	1.10	0.10	0.18	0.11	1.87	1.68	2.55
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales	1.60	1.91	2.85	0.06	2.31	0.45	1.79
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales	8.20	2.71	4.02	0.61	17.99	6.60	17.41
	k__Bacteria;p__Armatimonadetes;c__0319-6E2;o__	0.00	0.01	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Armatimonadetes;c__Armatimonadia;o__Armatimonadales	0.00	0.00	0.01	0.00	0.03	0.01	0.02
	k__Bacteria;p__Armatimonadetes;c__[Fimbriimonadia];o__[Fimbriimonadales]	0.00	0.11	0.07	0.00	0.01	0.00	0.00
	k__Bacteria;p__BRC1;c__PRR-11;o__	0.00	0.01	0.00	0.03	0.08	0.07	0.01
	k__Bacteria;p__Bacteroidetes;c__At12OctB3;o__	0.00	0.00	0.01	0.00	0.11	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales	3.30	3.57	1.86	8.22	0.52	3.97	1.53
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales	2.20	0.17	0.47	12.70	0.04	0.02	0.03

	Orders	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales	1.20	0.94	2.03	0.29	0.29	3.33	0.16
	k__Bacteria;p__Bacteroidetes;c__[Rhodothermi];o__[Rhodothermales]	0.10	0.13	0.07	0.43	0.01	0.01	0.02
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales]	4.40	5.37	5.46	1.10	4.12	5.38	5.24
	k__Bacteria;p__Chlamydiae;c__Chlamydiia;o__	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chlorobi;c__OPB56;o__	0.10	0.20	0.09	0.03	0.19	0.00	0.02
	k__Bacteria;p__Chlorobi;c__SJA-28;o__	0.10	0.45	0.25	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__;o__	0.00	0.01	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__	0.00	0.02	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__A31	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Ardenscatenales	0.10	0.18	0.20	0.03	0.09	0.00	0.04
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Caldilineales	0.00	0.11	0.10	0.02	0.01	0.00	0.03
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__H39	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__SBR1031	0.10	0.39	0.12	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__WCHB1-50	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__envOPS12	0.10	0.28	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__C0119;o__	0.00	0.09	0.03	0.00	0.10	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__AKIW781	0.00	0.01	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Chloroflexales	0.00	0.26	0.01	0.00	0.01	0.01	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Herpetosiphonales	0.00	0.03	0.01	0.00	0.01	0.00	0.01
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__[Roseiflexales]	0.00	0.13	0.05	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Ellin6529;o__	0.70	1.00	1.05	0.15	0.82	0.26	0.62
	k__Bacteria;p__Chloroflexi;c__Gitt-GS-136;o__	0.60	0.30	0.68	0.13	1.09	0.50	0.83
	k__Bacteria;p__Chloroflexi;c__P2-11E;o__	0.00	0.04	0.03	0.00	0.03	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__S085;o__	0.10	0.38	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__TK10;Other	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Chloroflexi;c__TK10;o__	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Chloroflexi;c__TK10;o__AKYG885	0.10	0.06	0.05	0.00	0.18	0.02	0.21
	k__Bacteria;p__Chloroflexi;c__TK17;o__mle1-48	0.00	0.09	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia;o__AKYG1722	0.10	0.00	0.17	0.05	0.03	0.16	0.08
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia;o__JG30-KF-CM45	1.20	0.73	1.60	0.42	1.31	1.70	1.74

	Orders	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Chloroflexi;c__[Thermobacula];o__[Thermobaculales]	0.00	0.00	0.00	0.00	0.02	0.02	0.03
	k__Bacteria;p__Cyanobacteria;c__;o__	2.60	15.50	0.09	0.00	0.09	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Chlorophyta	0.00	0.05	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Stramenopiles	0.50	0.22	1.48	0.37	0.64	0.09	0.07
	k__Bacteria;p__Cyanobacteria;c__ML635J-21;o__	0.00	0.04	0.03	0.00	0.03	0.08	0.02
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae;Other	0.00	0.04	0.00	0.00	0.03	0.01	0.03
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae;o__Nostocales	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae;o__Chroococcales	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae;o__Oscillatoriales	2.00	1.19	10.19	0.83	0.01	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales	2.70	4.89	1.79	9.51	0.03	0.02	0.06
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Synechococcales	0.00	0.12	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__FAC88	0.00	0.06	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__IIb	0.00	0.14	0.07	0.03	0.03	0.00	0.02
	k__Bacteria;p__FBP;c__;o__	0.00	0.01	0.05	0.00	0.00	0.06	0.02
	k__Bacteria;p__Fibrobacteres;c__Fibrobacteria;o__258ds10	0.00	0.02	0.01	0.01	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales	0.10	0.01	0.36	0.03	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales	0.00	0.03	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__BB34;o__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__GKS2-174;o__	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__;o__	0.10	0.05	0.82	0.00	0.02	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-1;o__	0.40	0.58	0.94	0.13	0.37	0.01	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-3;o__	0.70	0.00	0.53	0.79	0.91	0.84	0.83
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-5;o__	0.10	0.11	0.17	0.05	0.08	0.01	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;Other	0.10	0.15	0.12	0.01	0.11	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__	0.30	0.28	0.91	0.04	0.30	0.00	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Ellin5290	0.00	0.06	0.17	0.00	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales	1.00	1.61	0.69	0.31	0.61	2.38	0.49
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__KD8-87	0.10	0.08	0.09	0.00	0.31	0.00	0.04

	Orders	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__N1423WL	0.60	0.78	1.34	0.05	1.02	0.02	0.38
	k__Bacteria;p__Nitrospirae;c__Nitrospira;o__Nitrospirales	0.10	0.60	0.13	0.02	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__;o__	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__Mb-NB09;o__	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__OD1;c__SM2F11;o__	0.00	0.00	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__OP11;c__WCHB1-64;o__d153	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Phycisphaerae;o__Phycisphaerales	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Gemmatales	0.00	0.03	0.00	0.00	0.02	0.00	0.04
	k__Bacteria;p__Planctomycetes;c__vadinHA49;o__DH61	0.00	0.02	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__	0.00	0.12	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__BD7-3	0.00	0.09	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales	0.20	0.55	0.51	0.03	0.14	0.13	0.05
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales	2.00	2.90	2.16	1.56	2.36	0.61	2.20
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales	1.70	1.55	2.13	5.41	0.43	0.41	0.42
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales	0.30	1.13	0.22	0.13	0.08	0.16	0.22
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales	0.10	0.01	0.12	0.01	0.16	0.00	0.05
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales	8.60	9.63	8.61	4.78	10.88	9.07	8.37
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;Other	0.00	0.11	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__	0.10	0.11	0.05	0.02	0.10	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__A21b	0.00	0.10	0.14	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales	1.80	3.94	2.90	0.61	0.85	1.66	0.82
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Ellin6067	0.00	0.22	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__MND1	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Nitrosomonadales	0.00	0.11	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Procabacteriales	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__SC-I-84	0.20	0.51	0.43	0.03	0.03	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__	0.20	0.09	0.39	0.11	0.27	0.14	0.08
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales	0.40	0.25	0.48	0.29	0.47	0.72	0.48
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__FAC87	0.00	0.03	0.00	0.00	0.04	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__GMD14H09	0.00	0.00	0.00	0.09	0.00	0.00	0.00

	Orders	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__MIZ46	0.10	0.05	0.21	0.02	0.16	0.10	0.08
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales	1.00	1.60	1.43	1.11	0.82	0.57	0.47
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__PB19	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Spirobacillales	0.20	0.32	0.18	0.13	0.13	0.09	0.15
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Syntrophobacterales	0.00	0.01	0.09	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__	0.00	0.00	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales	0.10	0.00	0.05	0.28	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Chromatiales	0.10	0.01	0.20	0.00	0.10	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__HTCC2188	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales	0.10	0.19	0.40	0.06	0.15	0.02	0.05
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales	0.00	0.00	0.03	0.00	0.00	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales	0.00	0.01	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Thiotrichales	0.20	0.23	0.68	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales	1.40	3.05	2.43	2.52	0.27	0.05	0.03
	k__Bacteria;p__Proteobacteria;c__TA18;o__CV90	0.00	0.14	0.01	0.00	0.03	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__TA18;o__PHOS-HD29	0.10	0.19	0.20	0.00	0.15	0.00	0.04
	k__Bacteria;p__TM6;c__SJA-4;o__	0.00	0.00	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__TM7;c__SC3;o__	0.00	0.00	0.12	0.03	0.05	0.00	0.02
	k__Bacteria;p__TM7;c__TM7-1;o__	1.10	0.30	1.98	0.71	2.44	0.29	0.61
	k__Bacteria;p__TM7;c__TM7-3;o__	0.00	0.00	0.01	0.02	0.00	0.03	0.00
	k__Bacteria;p__TM7;c__TM7-3;o__I025	0.10	0.00	0.07	0.26	0.03	0.07	0.03
	k__Bacteria;p__Verrucomicrobia;c__Opitutae;o__Opitutales	0.00	0.05	0.07	0.02	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales	0.10	0.27	0.25	0.34	0.01	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Methylacidiphilae];o__S-BQ2-57	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales]	0.10	0.12	0.25	0.05	0.11	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales]	1.50	1.01	1.50	0.17	2.66	0.13	3.32
	k__Bacteria;p__WPS-2;c__;o__	0.00	0.05	0.00	0.00	0.01	0.00	0.01
	k__Bacteria;p__WS3;c__PRR-12;o__Sediment-1	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales	0.40	0.05	0.10	1.38	0.09	0.55	0.12

Appendix 4. Family-based taxonomy on QIIME

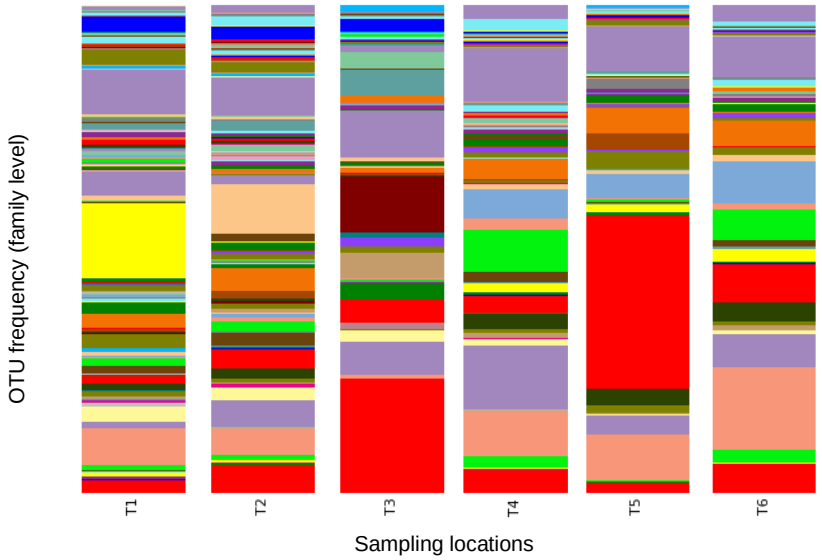


Figure 1. Relative abundance of family-affiliated OTUs per sampling point. Color code is presented in Table 1.

Table 1. Frequency of OTU abundance at the family-level. All values are percentages.

	Families	Total	T1	T2	T3	T4	T5	T6
	Unassigned;Other;Other;Other;Other	7.40	2.71	5.77	23.34	4.76	2.19	5.91
	k__Bacteria;p__Acidobacteria;c__AT-s54;o__;f__	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__CCU21;f__	0.00	0.06	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__iii1-15;f__	0.10	0.14	0.10	0.08	0.03	0.09	0.04
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__iii1-15;f__mb2424	0.20	0.31	0.47	0.12	0.15	0.02	0.15
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__	0.30	0.97	0.30	0.00	0.30	0.01	0.14
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__Solibacteraceae	0.00	0.14	0.16	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__[Bryobacteraceae]	0.00	0.15	0.01	0.00	0.00	0.00	0.02
	k__Bacteria;p__Acidobacteria;c__Sva0725;o__Sva0725;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__DS-100;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__PK29;f__	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__RB41;f__	1.20	1.10	0.87	0.00	2.37	0.22	2.56
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__RB41;f__Ellin6075	8.30	7.54	5.49	0.73	9.38	9.53	16.87
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__32-20;f__	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__DS-18;f__	0.10	0.11	0.30	0.03	0.09	0.00	0.03
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__	6.20	1.29	5.40	6.63	13.15	3.94	6.84
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__C111	1.70	2.97	2.68	2.45	1.20	0.23	0.81
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__EB1017	0.20	0.93	0.04	0.00	0.02	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__Iamiaceae	0.30	0.56	0.75	0.00	0.38	0.01	0.03
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__Microthrixaceae	0.00	0.08	0.03	0.00	0.00	0.01	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;Other	0.60	0.64	0.30	0.00	1.05	0.27	1.10
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__	0.80	1.04	0.65	0.30	0.69	1.32	0.66
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Geodermatophilaceae	0.00	0.00	0.00	0.00	0.03	0.01	0.02
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Intrasporangiaceae	0.20	0.04	0.08	1.18	0.00	0.00	0.02
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae	0.00	0.10	0.09	0.00	0.02	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micromonosporaceae	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardioidaceae	2.30	1.35	1.86	0.13	3.24	3.56	3.80

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Pseudonocardiaceae	0.00	0.01	0.01	0.01	0.03	0.02	0.03
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Sporichthyaceae	9.50	1.64	3.94	4.63	3.45	35.39	7.89
	k__Bacteria;p__Actinobacteria;c__MB-A2-108;o__;f__	0.00	0.00	0.00	0.00	0.00	0.00	0.02
	k__Bacteria;p__Actinobacteria;c__MB-A2-108;o__0319-7L14;f__	0.10	0.10	0.34	0.00	0.23	0.01	0.15
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria;o__Euzebyales;f__Euzebyaceae	0.80	0.03	0.16	3.32	0.49	0.62	0.36
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria;o__Nitriliruptorales;f__Nitriliruptoraceae	0.00	0.00	0.00	0.28	0.00	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Rubrobacteria;o__Rubrobacterales;f__Rubrobacteraceae	1.10	0.10	0.18	0.11	1.87	1.68	2.55
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;Other	0.00	0.01	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;f__	0.10	0.20	0.03	0.00	0.04	0.09	0.11
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;f__AK1AB1_02E	0.20	0.03	0.09	0.06	0.27	0.26	0.44
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;f__Gaiellaceae	1.30	1.66	2.73	0.00	1.99	0.11	1.24
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;Other	0.00	0.00	0.13	0.00	0.14	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__	3.30	1.42	2.12	0.36	8.56	0.68	6.37
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Conexibacteraceae	0.80	0.20	0.70	0.18	2.43	0.05	1.33
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Patulibacteraceae	3.50	0.45	0.79	0.02	5.92	4.98	8.55
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Solirubrobacteraceae	0.70	0.63	0.27	0.06	0.95	0.90	1.15
	k__Bacteria;p__Armatimonadetes;c__0319-6E2;o__;f__	0.00	0.01	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Armatimonadetes;c__Armatimonadia;o__Armatimonadales;f__Armatimonadaceae	0.00	0.00	0.01	0.00	0.03	0.01	0.02
	k__Bacteria;p__Armatimonadetes;c__[Fimbriimonadia];o__[Fimbriimonadales];f__[Fimbriimonadaceae]	0.00	0.11	0.07	0.00	0.01	0.00	0.00
	k__Bacteria;p__BRC1;c__PRR-11;o__;f__	0.00	0.01	0.00	0.03	0.08	0.07	0.01
	k__Bacteria;p__Bacteroidetes;c__At12OctB3;o__;f__	0.00	0.00	0.01	0.00	0.11	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__	0.10	0.70	0.00	0.00	0.00	0.09	0.05
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cyclobacteriaceae	1.00	0.04	0.66	5.05	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae	1.80	2.81	1.03	1.37	0.52	3.80	1.48
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Flammeovirgaceae	0.30	0.01	0.17	1.79	0.00	0.08	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__[Amoebophilaceae]	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae	0.20	0.13	0.01	1.07	0.04	0.00	0.03
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae	2.00	0.04	0.46	11.64	0.00	0.02	0.00
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__	0.20	0.38	0.47	0.26	0.06	0.02	0.03
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__Sphingobacteriaceae	1.00	0.56	1.56	0.03	0.23	3.31	0.13

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Bacteroidetes;c__[Rhodothermi];o__[Rhodothermales];f__Rhodothermaceae	0.10	0.13	0.07	0.43	0.01	0.01	0.02
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae	3.90	3.04	4.58	1.09	4.12	5.38	5.22
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Saprospiraceae	0.50	2.32	0.87	0.01	0.00	0.00	0.02
	k__Bacteria;p__Chlamydiae;c__Chlamydiia;o__;f__	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chlorobi;c__OPB56;o__;f__	0.10	0.20	0.09	0.03	0.19	0.00	0.02
	k__Bacteria;p__Chlorobi;c__SJA-28;o__;f__	0.10	0.45	0.25	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__;o__;f__	0.00	0.01	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__;f__	0.00	0.02	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__A31;f__S47	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Ardenscatenales;f__Ardenscatenaceae	0.10	0.18	0.20	0.03	0.09	0.00	0.04
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Caldilineales;f__Caldilineaceae	0.00	0.11	0.10	0.02	0.01	0.00	0.03
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__H39;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__SBR1031;f__A4b	0.10	0.39	0.12	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__WCHB1-50;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__envOPS12;f__	0.10	0.28	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__C0119;o__;f__	0.00	0.09	0.03	0.00	0.10	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__AKIW781;f__	0.00	0.01	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Chloroflexales;Other	0.00	0.12	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Chloroflexales;f__Chloroflexaceae	0.00	0.14	0.01	0.00	0.00	0.01	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Herpetosiphonales;f__	0.00	0.03	0.01	0.00	0.01	0.00	0.01
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__[Roseiflexales];f__[Kouleothrixaceae]	0.00	0.13	0.05	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Ellin6529;o__;f__	0.70	1.00	1.05	0.15	0.82	0.26	0.62
	k__Bacteria;p__Chloroflexi;c__Gitt-GS-136;o__;f__	0.60	0.30	0.68	0.13	1.09	0.50	0.83
	k__Bacteria;p__Chloroflexi;c__P2-11E;o__;f__	0.00	0.04	0.03	0.00	0.03	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__S085;o__;f__	0.10	0.38	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__TK10;Other;Other	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Chloroflexi;c__TK10;o__;f__	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Chloroflexi;c__TK10;o__AKYG885;f__	0.10	0.02	0.04	0.00	0.18	0.02	0.21
	k__Bacteria;p__Chloroflexi;c__TK10;o__AKYG885;f__Dolo_23	0.00	0.04	0.01	0.00	0.00	0.00	0.00

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Chloroflexi;c__TK17;o__mle1-48;f__	0.00	0.09	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia;o__AKYG1722;f__	0.10	0.00	0.17	0.05	0.03	0.16	0.08
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia;o__JG30-KF-CM45;f__	1.20	0.73	1.60	0.42	1.31	1.70	1.74
	k__Bacteria;p__Chloroflexi;c__[Thermobacula];o__[Thermobaculales];f__[Thermobaculaceae]	0.00	0.00	0.00	0.00	0.02	0.02	0.03
	k__Bacteria;p__Cyanobacteria;c__;o__;f__	2.60	15.50	0.09	0.00	0.09	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__;f__	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Chlorophyta;f__	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Chlorophyta;f__Chlamydomonadaceae	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Stramenopiles;f__	0.50	0.22	1.48	0.37	0.64	0.09	0.07
	k__Bacteria;p__Cyanobacteria;c__ML635J-21;o__;f__	0.00	0.04	0.03	0.00	0.03	0.08	0.02
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae;Other;Other	0.00	0.04	0.00	0.00	0.03	0.01	0.03
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae;o__Nostocales;f__Nostocaceae	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae;o__Chroococcales;f__Xenococcaceae	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae;o__Oscillatoriales;f__Phormidiaceae	2.00	1.19	10.19	0.83	0.01	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales;f__Pseudanabaenaceae	2.70	4.87	1.79	9.51	0.03	0.02	0.06
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Synechococcales;f__Chamaesiphonaceae	0.00	0.12	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__FAC88;f__	0.00	0.06	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__IIb;f__	0.00	0.14	0.07	0.03	0.03	0.00	0.02
	k__Bacteria;p__FBP;c__;o__;f__	0.00	0.01	0.05	0.00	0.00	0.06	0.02
	k__Bacteria;p__Fibrobacteres;c__Fibrobacteria;o__258ds10;f__	0.00	0.02	0.01	0.01	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae	0.00	0.01	0.09	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Paenibacillaceae	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae	0.00	0.00	0.27	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__BB34;o__;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__GKS2-174;o__;f__	0.00	0.04	0.00	0.00	0.00	0.00	0.00

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Gemmatimonadetes;c__;o__;f__	0.10	0.05	0.82	0.00	0.02	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-1;o__;f__	0.40	0.58	0.94	0.13	0.37	0.01	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-3;o__;f__	0.70	0.00	0.53	0.79	0.91	0.84	0.83
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-5;o__;f__	0.10	0.11	0.17	0.05	0.08	0.01	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;Other;Other	0.10	0.15	0.12	0.01	0.11	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__;f__	0.30	0.28	0.91	0.04	0.30	0.00	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Ellin5290;f__	0.00	0.06	0.17	0.00	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;Other	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__	0.50	0.17	0.07	0.00	0.23	1.99	0.32
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__A1-B1	0.10	0.83	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__Ellin5301	0.20	0.31	0.34	0.01	0.32	0.33	0.16
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__Gemmatimonadaceae	0.20	0.29	0.29	0.27	0.07	0.05	0.01
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__KD8-87;f__	0.10	0.08	0.09	0.00	0.31	0.00	0.04
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__N1423WL;f__	0.60	0.78	1.34	0.05	1.02	0.02	0.38
	k__Bacteria;p__Nitrospirae;c__Nitrospira;o__Nitrospirales;f__Nitrospiraceae	0.10	0.60	0.13	0.02	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__;o__;f__	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__Mb-NB09;o__;f__	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__OD1;c__SM2F11;o__;f__	0.00	0.00	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__OP11;c__WCHB1-64;o__d153;f__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Phycisphaerae;o__Phycisphaerales;f__	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Gemmatales;f__Gemmataceae	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Gemmatales;f__Isosphaeraceae	0.00	0.01	0.00	0.00	0.02	0.00	0.04
	k__Bacteria;p__Planctomycetes;c__vadinHA49;o__DH61;f__	0.00	0.02	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__;f__	0.00	0.12	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__BD7-3;f__	0.00	0.09	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae	0.20	0.55	0.51	0.03	0.14	0.13	0.05
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;Other	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__	0.30	0.82	0.70	0.03	0.23	0.00	0.07
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Aurantimonadaceae	0.00	0.00	0.20	0.09	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae	0.50	0.52	0.09	1.44	0.49	0.10	0.62

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae	0.10	0.18	0.20	0.00	0.12	0.08	0.16
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae	0.20	0.85	0.38	0.00	0.08	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Phyllobacteriaceae	0.10	0.19	0.17	0.00	0.17	0.04	0.05
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhodobiaceae	0.60	0.31	0.43	0.00	1.27	0.40	1.30
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Hyphomonadaceae	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae	1.70	1.54	2.13	5.41	0.43	0.41	0.42
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__	0.00	0.03	0.00	0.07	0.04	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae	0.30	0.91	0.21	0.06	0.03	0.16	0.22
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae	0.00	0.19	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__	0.00	0.01	0.07	0.01	0.13	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__Rickettsiaceae	0.00	0.00	0.05	0.00	0.03	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__	0.20	0.54	0.46	0.00	0.00	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Erythrobacteraceae	0.60	0.01	0.23	3.37	0.00	0.01	0.01
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae	7.80	9.08	7.92	1.41	10.88	9.06	8.34
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;Other;Other	0.00	0.11	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__;f__	0.10	0.11	0.05	0.02	0.10	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__A21b;f__UD5	0.00	0.10	0.14	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae	0.20	0.46	0.42	0.05	0.09	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae	0.20	0.25	0.23	0.11	0.15	0.29	0.20
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae	1.30	3.17	2.20	0.45	0.34	1.05	0.38
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae	0.20	0.05	0.05	0.01	0.28	0.32	0.22
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Ellin6067;f__	0.00	0.22	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__MND1;f__	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Nitrosomonadales;f__Nitrosomonadaceae	0.00	0.11	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Procabacteriales;f__Procabacteriaceae	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__SC-I-84;f__	0.20	0.51	0.43	0.03	0.03	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__;f__	0.20	0.09	0.39	0.11	0.27	0.14	0.08
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales;f__Bacteriovoracaceae	0.30	0.17	0.30	0.24	0.30	0.51	0.38
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales;f__Bdellovibrionaceae	0.10	0.08	0.18	0.04	0.17	0.21	0.10
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__FAC87;f__	0.00	0.03	0.00	0.00	0.04	0.00	0.01

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__GMD14H09;f__	0.00	0.00	0.00	0.09	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__MIZ46;f__	0.10	0.05	0.21	0.02	0.16	0.10	0.08
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__	0.60	1.23	0.77	0.45	0.40	0.47	0.30
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__0319-6G20	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Cystobacterineae	0.00	0.01	0.03	0.00	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Haliangiaceae	0.20	0.26	0.36	0.06	0.21	0.06	0.09
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Myxococcaceae	0.00	0.01	0.10	0.01	0.01	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Nannocystaceae	0.10	0.06	0.01	0.60	0.04	0.00	0.05
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__OM27	0.00	0.02	0.16	0.00	0.10	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Polyangiaceae	0.00	0.00	0.00	0.00	0.05	0.04	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__PB19;f__	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Spirobacillales;f__	0.20	0.32	0.18	0.13	0.13	0.09	0.15
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Syntrophobacterales;f__Syntrophaceae	0.00	0.00	0.07	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Syntrophobacterales;f__Syntrophobacteraceae	0.00	0.01	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__;f__	0.00	0.00	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__Alteromonadaceae	0.00	0.00	0.00	0.14	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__HTCC2188	0.00	0.00	0.05	0.13	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Chromatiales;f__	0.10	0.01	0.20	0.00	0.10	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__HTCC2188;f__HTCC2089	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__	0.00	0.09	0.10	0.03	0.01	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Coxiellaceae	0.10	0.06	0.29	0.01	0.10	0.02	0.02
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Legionellaceae	0.00	0.04	0.01	0.02	0.04	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Hahellaceae	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Halomonadaceae	0.00	0.00	0.00	0.00	0.00	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae	0.00	0.01	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Thiotrichales;f__Piscirickettsiaceae	0.20	0.23	0.68	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae	1.40	3.05	2.43	2.52	0.27	0.05	0.03
	k__Bacteria;p__Proteobacteria;c__TA18;o__CV90;f__	0.00	0.14	0.01	0.00	0.03	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__TA18;o__PHOS-HD29;f__	0.10	0.19	0.20	0.00	0.15	0.00	0.04
	k__Bacteria;p__TM6;c__SJA-4;o__;f__	0.00	0.00	0.01	0.00	0.01	0.00	0.00

	Families	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__TM7;c__SC3;o__;f__	0.00	0.00	0.12	0.03	0.05	0.00	0.02
	k__Bacteria;p__TM7;c__TM7-1;o__;f__	1.10	0.30	1.98	0.71	2.44	0.29	0.61
	k__Bacteria;p__TM7;c__TM7-3;o__;f__	0.00	0.00	0.01	0.02	0.00	0.03	0.00
	k__Bacteria;p__TM7;c__TM7-3;o__I025;f__	0.10	0.00	0.07	0.26	0.03	0.07	0.03
	k__Bacteria;p__Verrucomicrobia;c__Opitutae;o__Opitutales;f__Opitutaceae	0.00	0.05	0.07	0.02	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae	0.10	0.27	0.25	0.34	0.01	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Methylacidiphilae];o__S-BQ2-57;f__	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__	0.10	0.08	0.21	0.04	0.11	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__Ellin515	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__Ellin517	0.00	0.01	0.04	0.01	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__R4-41B	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae]	1.50	1.01	1.50	0.17	2.66	0.13	3.32
	k__Bacteria;p__WPS-2;c__;o__;f__	0.00	0.05	0.00	0.00	0.01	0.00	0.01
	k__Bacteria;p__WS3;c__PRR-12;o__Sediment-1;f__PRR-10	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Deinococcaceae	0.00	0.04	0.00	0.00	0.03	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Trueperaceae	0.40	0.01	0.10	1.38	0.06	0.55	0.12

Appendix 5. Genera-based taxonomy on QIIME

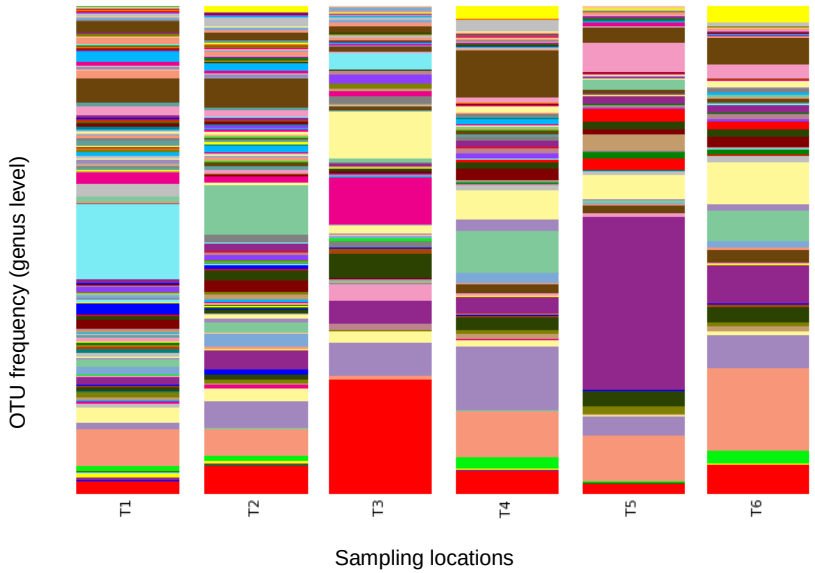


Figure 1. Relative abundance of genus-affiliated OTUs per sampling point. Color code is presented in Table 1.

Table 1. Frequency of OTU abundance at the genus-level. All values are percentages.

	Genera	Total	T1	T2	T3	T4	T5	T6
	Unassigned;Other;Other;Other;Other;Other	7.40	2.71	5.77	23.34	4.76	2.19	5.91
	k__Bacteria;p__Acidobacteria;c__AT-s54;o__;f__;g__	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__CCU21;f__;g__	0.00	0.06	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__iii1-15;f__;g__	0.10	0.14	0.10	0.08	0.03	0.09	0.04
	k__Bacteria;p__Acidobacteria;c__Acidobacteria-6;o__iii1-15;f__mb2424;g__	0.20	0.31	0.47	0.12	0.15	0.02	0.15
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__;g__	0.30	0.97	0.30	0.00	0.30	0.01	0.14
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__Solibacteraceae;g__Candidatus	0.00	0.16	0.00	0.00	0.00	0.14	0.00
	k__Bacteria;p__Acidobacteria;c__Solibacteres;o__Solibacterales;f__[Bryobacteraceae];g__	0.00	0.15	0.01	0.00	0.00	0.00	0.02
	k__Bacteria;p__Acidobacteria;c__Sva0725;o__Sva0725;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__DS-100;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__PK29;f__;g__	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__RB41;f__;g__	1.20	1.10	0.87	0.00	2.37	0.22	2.56
	k__Bacteria;p__Acidobacteria;c__[Chloracidobacteria];o__RB41;f__Ellin6075;g__	8.30	7.54	5.49	0.73	9.38	9.53	16.87
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__32-20;f__;g__	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Acidobacteria;c__iii1-8;o__DS-18;f__;g__	0.10	0.11	0.30	0.03	0.09	0.00	0.03
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__;g__	6.20	1.29	5.40	6.63	13.15	3.94	6.84
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__C111;g__	1.70	2.97	2.68	2.45	1.20	0.23	0.81
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__EB1017;g__	0.20	0.93	0.04	0.00	0.02	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__Iamiaceae;g__Iamia	0.30	0.56	0.75	0.00	0.38	0.01	0.03
	k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__Microthrixaceae;g__	0.00	0.08	0.03	0.00	0.00	0.01	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;Other;Other	0.60	0.64	0.30	0.00	1.05	0.27	1.10
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__;g__	0.80	1.04	0.65	0.30	0.69	1.32	0.66
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Geodermatophilaceae;Other	0.00	0.00	0.00	0.00	0.03	0.01	0.02
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Intrasporangiaceae;g__	0.20	0.04	0.08	1.18	0.00	0.00	0.02

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Arthrobacter	0.00	0.10	0.09	0.00	0.02	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micromonosporaceae;g__	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardioidaceae;g__	1.80	0.96	0.96	0.01	2.48	3.22	3.17
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardioidaceae;g__Aeromicrobium	0.20	0.10	0.01	0.00	0.44	0.01	0.39
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardioidaceae;g__Friedmanniella	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Nocardioidaceae;g__Nocardioides	0.40	0.29	0.88	0.13	0.32	0.32	0.24
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Pseudonocardiaceae;g__	0.00	0.01	0.00	0.00	0.03	0.02	0.03
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Pseudonocardiaceae;g__Pseudonocardia	0.00	0.00	0.01	0.01	0.00	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Sporichthyaceae;g__	9.50	1.64	3.94	4.63	3.45	35.39	7.89
	k__Bacteria;p__Actinobacteria;c__MB-A2-108;o__;f__;g__	0.00	0.00	0.00	0.00	0.00	0.00	0.02
	k__Bacteria;p__Actinobacteria;c__MB-A2-108;o__0319-7L14;f__;g__	0.10	0.10	0.34	0.00	0.23	0.01	0.15
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria;o__Euzebyales;f__Euzebyaceae;g__Euzebya	0.80	0.03	0.16	3.32	0.49	0.62	0.36
	k__Bacteria;p__Actinobacteria;c__Nitriliruptoria;o__Nitriliruptorales;f__Nitriliruptoraceae;g__	0.00	0.00	0.00	0.28	0.00	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Rubrobacteria;o__Rubrobacterales;f__Rubrobacteraceae;g__Rubrobacter	1.10	0.10	0.18	0.11	1.87	1.68	2.55
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;Other;Other	0.00	0.01	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;f__;g__	0.10	0.20	0.03	0.00	0.04	0.09	0.11
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;f__AK1AB1_02E;g__	0.20	0.03	0.09	0.06	0.27	0.26	0.44
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Gaiellales;f__Gaiellaceae;g__	1.30	1.66	2.73	0.00	1.99	0.11	1.24
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;Other;Other	0.00	0.00	0.13	0.00	0.14	0.00	0.00
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__;g__	3.30	1.42	2.12	0.36	8.56	0.68	6.37
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Conexibacteraceae;g__	0.80	0.20	0.70	0.18	2.43	0.05	1.33
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Patulibacteraceae;g__	3.50	0.45	0.79	0.02	5.92	4.98	8.55
	k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Solirubrobacteraceae;g__	0.70	0.63	0.27	0.06	0.95	0.90	1.15
	k__Bacteria;p__Armatimonadetes;c__0319-6E2;o__;f__;g__	0.00	0.01	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Armatimonadetes;c__Armatimonadia;o__Armatimonadales;f__Armatimonadaceae;g__	0.00	0.00	0.01	0.00	0.03	0.01	0.02
	k__Bacteria;p__Armatimonadetes;c__[Fimbriimonadia];o__[Fimbriimonadales];f__[Fimbriimonadaceae];g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Armatimonadetes;c__[Fimbriimonadia];o__[Fimbriimonadales];f__[Fimbriimonadaceae];g__Fimbriimonas	0.00	0.10	0.07	0.00	0.01	0.00	0.00
	k__Bacteria;p__BRC1;c__PRR-11;o__;f__;g__	0.00	0.01	0.00	0.03	0.08	0.07	0.01
	k__Bacteria;p__Bacteroidetes;c__At12OctB3;o__;f__;g__	0.00	0.00	0.01	0.00	0.11	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__;g__	0.10	0.70	0.00	0.00	0.00	0.09	0.05

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cyclobacteriaceae;Other	0.00	0.00	0.00	0.10	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cyclobacteriaceae;g__	0.90	0.04	0.66	4.95	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__	0.30	0.63	0.21	1.08	0.02	0.08	0.03
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Adhaeribacter	0.40	0.00	0.03	0.15	0.03	2.28	0.16
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Dyadobacter	0.00	0.00	0.18	0.09	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Flectobacillus	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Hymenobacter	0.60	0.71	0.18	0.00	0.39	1.30	1.19
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Persicitalea	0.00	0.00	0.01	0.02	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Pontibacter	0.10	0.04	0.25	0.00	0.00	0.05	0.02
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Rhodocytophaga	0.00	0.02	0.01	0.00	0.01	0.03	0.02
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Rudanella	0.10	0.81	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Spirosoma	0.10	0.47	0.01	0.00	0.03	0.02	0.06
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Sporocytophaga	0.10	0.08	0.14	0.03	0.03	0.05	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Flammeovirgaceae;g__	0.30	0.00	0.17	1.45	0.00	0.08	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Flammeovirgaceae;g__Flexibacter	0.10	0.01	0.00	0.34	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__[Amoebophilaceae];g__	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae;g__	0.20	0.12	0.01	0.71	0.04	0.00	0.03
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae;g__Crocinitomix	0.00	0.01	0.00	0.08	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae;g__Cryomorpha	0.00	0.00	0.00	0.18	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Cryomorphaceae;g__Fluviicola	0.00	0.00	0.00	0.08	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae;g__	0.30	0.00	0.23	1.74	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae;g__Flavobacterium	0.00	0.04	0.01	0.10	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae;g__Gillisia	1.70	0.00	0.21	9.80	0.00	0.02	0.00
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__;g__	0.20	0.38	0.47	0.26	0.06	0.02	0.03
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__Sphingobacteriaceae;g__	0.90	0.51	1.04	0.02	0.23	3.31	0.13
	k__Bacteria;p__Bacteroidetes;c__Sphingobacteriia;o__Sphingobacteriales;f__Sphingobacteriaceae;g__Pedobacter	0.10	0.05	0.52	0.01	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__[Rhodothermi];o__[Rhodothermales];f__Rhodothermaceae;g__	0.00	0.00	0.00	0.24	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__[Rhodothermi];o__[Rhodothermales];f__Rhodothermaceae;g__Rubricoccus	0.10	0.13	0.07	0.18	0.01	0.01	0.02
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__;g__	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__	1.80	1.91	2.46	0.52	2.48	1.23	2.08

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__Flavisolibacter	1.20	0.72	1.91	0.52	1.20	1.41	1.57
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__Sediminibacterium	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__Segetibacter	0.90	0.40	0.21	0.05	0.43	2.74	1.57
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Saprospiraceae;g__	0.50	2.31	0.87	0.00	0.00	0.00	0.02
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Saprospiraceae;g__Rubidimonas	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Saprospiraceae;g__Saprospira	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chlamydiae;c__Chlamydiia;o__;f__;g__	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chlorobi;c__OPB56;o__;f__;g__	0.10	0.20	0.09	0.03	0.19	0.00	0.02
	k__Bacteria;p__Chlorobi;c__SJA-28;o__;f__;g__	0.10	0.45	0.25	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__;o__;f__;g__	0.00	0.01	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__;f__;g__	0.00	0.02	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__A31;f__S47;g__	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Ardenscatenales;f__Ardenscatenaceae;g__Ardenscatena	0.10	0.18	0.20	0.03	0.09	0.00	0.04
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__Caldilineales;f__Caldilineaceae;g__	0.00	0.11	0.10	0.02	0.01	0.00	0.03
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__H39;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__SBR1031;f__A4b;g__	0.10	0.39	0.12	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__WCHB1-50;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__envOPS12;f__;g__	0.10	0.28	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__C0119;o__;f__;g__	0.00	0.09	0.03	0.00	0.10	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__AKIW781;f__;g__	0.00	0.01	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Chloroflexales;Other;Other	0.00	0.12	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Chloroflexales;f__Chloroflexaceae;Other	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Chloroflexales;f__Chloroflexaceae;g__Chloronema	0.00	0.11	0.01	0.00	0.00	0.01	0.00
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__Herpetosiphonales;f__;g__	0.00	0.03	0.01	0.00	0.01	0.00	0.01
	k__Bacteria;p__Chloroflexi;c__Chloroflexi;o__[Roseiflexales];f__[Kouleothrixaceae];g__	0.00	0.13	0.05	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Ellin6529;o__;f__;g__	0.70	1.00	1.05	0.15	0.82	0.26	0.62
	k__Bacteria;p__Chloroflexi;c__Gitt-GS-136;o__;f__;g__	0.60	0.30	0.68	0.13	1.09	0.50	0.83
	k__Bacteria;p__Chloroflexi;c__P2-11E;o__;f__;g__	0.00	0.04	0.03	0.00	0.03	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__S085;o__;f__;g__	0.10	0.38	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__TK10;Other;Other;Other	0.00	0.00	0.00	0.00	0.00	0.01	0.00

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Chloroflexi;c__TK10;o__;f__;g__	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Chloroflexi;c__TK10;o__AKYG885;f__;g__	0.10	0.02	0.04	0.00	0.18	0.02	0.21
	k__Bacteria;p__Chloroflexi;c__TK10;o__AKYG885;f__Dolo_23;g__	0.00	0.04	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__TK17;o__mle1-48;f__;g__	0.00	0.09	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia;o__AKYG1722;f__;g__	0.10	0.00	0.17	0.05	0.03	0.16	0.08
	k__Bacteria;p__Chloroflexi;c__Thermomicrobia;o__JG30-KF-CM45;f__;g__	1.20	0.73	1.60	0.42	1.31	1.70	1.74
	k__Bacteria;p__Chloroflexi;c__[Thermobacula];o__[Thermobaculales];f__[Thermobaculaceae];g__Thermobaculum	0.00	0.00	0.00	0.00	0.02	0.02	0.03
	k__Bacteria;p__Cyanobacteria;c__;o__;f__;g__	2.60	15.50	0.09	0.00	0.09	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__;f__;g__	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Chlorophyta;f__;g__	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Chlorophyta;f__Chlamydomonadaceae;Other	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Stramenopiles;f__;g__	0.50	0.22	1.48	0.37	0.64	0.09	0.07
	k__Bacteria;p__Cyanobacteria;c__ML635J-21;o__;f__;g__	0.00	0.04	0.03	0.00	0.03	0.08	0.02
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae;Other;Other;Other	0.00	0.04	0.00	0.00	0.03	0.01	0.03
	k__Bacteria;p__Cyanobacteria;c__Nostocophycideae;o__Nostocales;f__Nostocaceae;g__	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae;o__Chroococcales;f__Xenococcaceae;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Oscillatoriophycideae;o__Oscillatoriales;f__Phormidiaceae;g__Phormidium	2.00	1.19	10.19	0.83	0.01	0.02	0.01
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales;f__Pseudanabaenaceae;g__	1.70	0.04	0.56	9.51	0.01	0.00	0.00
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales;f__Pseudanabaenaceae;g__Leptolyngbya	0.40	2.44	0.13	0.00	0.01	0.00	0.04
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Pseudanabaenales;f__Pseudanabaenaceae;g__Pseudanabaena	0.60	2.39	1.11	0.00	0.01	0.02	0.02
	k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Synechococcales;f__Chamaesiphonaceae;g__	0.00	0.12	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__FAC88;f__;g__	0.00	0.06	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__Elusimicrobia;c__Elusimicrobia;o__IIb;f__;g__	0.00	0.14	0.07	0.03	0.03	0.00	0.02
	k__Bacteria;p__FBP;c__;o__;f__;g__	0.00	0.01	0.05	0.00	0.00	0.06	0.02
	k__Bacteria;p__Fibrobacteres;c__Fibrobacteria;o__258ds10;f__;g__	0.00	0.02	0.01	0.01	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Marinibacillus	0.00	0.01	0.08	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Paenibacillaceae;g__Paenibacillus	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae;g__Paenisporosarcina	0.00	0.00	0.27	0.00	0.00	0.00	0.00

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Desulfosporosinus	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;Other	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Pelosinus	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__BB34;o__;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__GN02;c__GKS2-174;o__;f__;g__	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__;o__;f__;g__	0.10	0.05	0.82	0.00	0.02	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-1;o__;f__;g__	0.40	0.58	0.94	0.13	0.37	0.01	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-3;o__;f__;g__	0.70	0.00	0.53	0.79	0.91	0.84	0.83
	k__Bacteria;p__Gemmatimonadetes;c__Gemm-5;o__;f__;g__	0.10	0.11	0.17	0.05	0.08	0.01	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;Other;Other;Other	0.10	0.15	0.12	0.01	0.11	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__;f__;g__	0.30	0.28	0.91	0.04	0.30	0.00	0.24
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Ellin5290;f__;g__	0.00	0.06	0.17	0.00	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;Other;Other	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__;g__	0.50	0.17	0.07	0.00	0.23	1.99	0.32
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__A1-B1;g__	0.10	0.83	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__Ellin5301;g__	0.20	0.31	0.34	0.01	0.32	0.33	0.16
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__Gemmatimonadaceae;g__Gemmatimonas	0.20	0.29	0.29	0.27	0.07	0.05	0.01
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__KD8-87;f__;g__	0.10	0.08	0.09	0.00	0.31	0.00	0.04
	k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__N1423WL;f__;g__	0.60	0.78	1.34	0.05	1.02	0.02	0.38
	k__Bacteria;p__Nitrospirae;c__Nitrospira;o__Nitrospirales;f__Nitrospiraceae;g__	0.00	0.00	0.08	0.02	0.00	0.00	0.00
	k__Bacteria;p__Nitrospirae;c__Nitrospira;o__Nitrospirales;f__Nitrospiraceae;g__Nitrospira	0.10	0.60	0.05	0.00	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__;o__;f__;g__	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__OD1;c__Mb-NB09;o__;f__;g__	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__OD1;c__SM2F11;o__;f__;g__	0.00	0.00	0.03	0.00	0.02	0.00	0.00
	k__Bacteria;p__OP11;c__WCHB1-64;o__d153;f__;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Phycisphaerae;o__Phycisphaerales;f__;g__	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Gemmatales;f__Gemmataceae;g__	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Planctomycetes;c__Planctomycetia;o__Gemmatales;f__Isosphaeraceae;g__	0.00	0.01	0.00	0.00	0.02	0.00	0.04

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Planctomycetes;c__vadinHA49;o__DH61;f__;g__	0.00	0.02	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__;f__;g__	0.00	0.12	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__BD7-3;f__;g__	0.00	0.09	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__	0.10	0.11	0.09	0.00	0.14	0.13	0.05
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Mycoplana	0.10	0.33	0.39	0.03	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Phenylobacterium	0.00	0.11	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;Other;Other	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__;g__	0.30	0.82	0.70	0.03	0.23	0.00	0.07
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Aurantimonadaceae;g__	0.00	0.00	0.20	0.09	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__	0.50	0.52	0.09	1.44	0.49	0.10	0.62
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae;g__	0.10	0.18	0.20	0.00	0.12	0.08	0.16
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__	0.10	0.54	0.30	0.00	0.04	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__Devosia	0.00	0.06	0.01	0.00	0.03	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__Rhodoplanes	0.10	0.25	0.07	0.00	0.01	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Phyllobacteriaceae;g__	0.00	0.13	0.05	0.00	0.01	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Phyllobacteriaceae;g__Mesorhizobium	0.10	0.06	0.12	0.00	0.16	0.03	0.05
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhodobiaceae;g__Afifella	0.60	0.31	0.43	0.00	1.27	0.40	1.30
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Hyphomonadaceae;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__	0.50	0.23	0.66	1.14	0.27	0.22	0.28
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Amaricoccus	0.00	0.14	0.08	0.03	0.00	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Anaerospora	0.10	0.00	0.05	0.50	0.00	0.02	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Jannaschia	0.20	0.00	0.04	1.05	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Loktanella	0.40	0.00	0.43	1.84	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Paracoccus	0.10	0.00	0.00	0.83	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Rhodobacter	0.10	0.49	0.33	0.01	0.01	0.01	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Rubellimicrobium	0.30	0.69	0.55	0.00	0.15	0.14	0.12
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__;g__	0.00	0.03	0.00	0.07	0.04	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__	0.20	0.63	0.21	0.03	0.03	0.14	0.19
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__Roseococcus	0.10	0.28	0.00	0.03	0.01	0.02	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__	0.00	0.19	0.01	0.00	0.00	0.00	0.00

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__;g__	0.00	0.01	0.07	0.01	0.13	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__Rickettsiaceae;g__	0.00	0.00	0.05	0.00	0.03	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__;g__	0.20	0.54	0.46	0.00	0.00	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Erythrobacteraceae;Other	0.00	0.00	0.03	0.00	0.00	0.01	0.01
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Erythrobacteraceae;g__	0.60	0.01	0.21	3.37	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__	2.20	1.96	1.42	0.31	1.03	5.88	2.63
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Blastomonas	0.20	0.60	0.29	0.00	0.02	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Kaistobacter	5.10	4.95	6.07	1.10	9.79	3.07	5.67
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Novosphingobium	0.00	0.03	0.07	0.00	0.02	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingomonas	0.00	0.00	0.07	0.00	0.01	0.08	0.03
	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Zymomonas	0.30	1.53	0.01	0.00	0.01	0.03	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;Other;Other;Other	0.00	0.11	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__;f__;g__	0.10	0.11	0.05	0.02	0.10	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__A21b;f__UD5;g__	0.00	0.10	0.14	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__	0.20	0.46	0.42	0.05	0.09	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Burkholderia	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Lautropia	0.20	0.25	0.23	0.11	0.15	0.28	0.20
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;Other	0.50	0.97	0.46	0.36	0.03	0.76	0.12
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__	0.70	1.85	1.60	0.08	0.31	0.29	0.25
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Hydrogenophaga	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Hylemonella	0.00	0.09	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Methylibium	0.10	0.22	0.12	0.01	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Ramlibacter	0.00	0.03	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__	0.00	0.04	0.04	0.01	0.04	0.13	0.03
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Janthinobacterium	0.10	0.01	0.01	0.00	0.24	0.19	0.19
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Ellin6067;f__;g__	0.00	0.22	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__MND1;f__;g__	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Nitrosomonadales;f__Nitrosomonadaceae;g__Nitrosovibrio	0.00	0.11	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Procabacteriales;f__Procabacteriaceae;g__Procabacter	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__SC-I-84;f__;g__	0.20	0.51	0.43	0.03	0.03	0.00	0.00

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__;f__;g__	0.20	0.09	0.39	0.11	0.27	0.14	0.08
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales;f__Bacteriovoracaceae;g__	0.30	0.17	0.27	0.24	0.29	0.51	0.38
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales;f__Bacteriovoracaceae;g__Bacteriovorax	0.00	0.00	0.03	0.00	0.01	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales;f__Bdellovibrionaceae;g__	0.00	0.02	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Bdellovibrionales;f__Bdellovibrionaceae;g__Bdellovibrio	0.10	0.05	0.16	0.04	0.17	0.21	0.10
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__FAC87;f__;g__	0.00	0.03	0.00	0.00	0.04	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__GMD14H09;f__;g__	0.00	0.00	0.00	0.09	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__MIZ46;f__;g__	0.10	0.05	0.21	0.02	0.16	0.10	0.08
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__;g__	0.60	1.23	0.77	0.45	0.40	0.47	0.30
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__0319-6G20;g__	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Cystobacterineae;g__	0.00	0.01	0.03	0.00	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Haliangiaceae;g__	0.20	0.26	0.36	0.06	0.21	0.06	0.09
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Myxococcaceae;g__	0.00	0.01	0.10	0.01	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Myxococcaceae;g__Anaeromyxobacter	0.00	0.00	0.00	0.00	0.01	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Nannocystaceae;g__Nannocystis	0.00	0.04	0.01	0.00	0.04	0.00	0.05
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Nannocystaceae;g__Plesiocystis	0.10	0.02	0.00	0.60	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__OM27;g__	0.00	0.02	0.16	0.00	0.10	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Polyangiaceae;Other	0.00	0.00	0.00	0.00	0.05	0.04	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__PB19;f__;g__	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Spirobacillales;f__;g__	0.20	0.32	0.18	0.13	0.13	0.09	0.15
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Syntrophobacterales;f__Syntrophaceae;g__Smithella	0.00	0.00	0.07	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Syntrophobacterales;f__Syntrophobacteraceae;g__	0.00	0.01	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__;f__;g__	0.00	0.00	0.00	0.00	0.01	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__Alteromonadaceae;g__Marinobacter	0.00	0.00	0.00	0.14	0.00	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__HTCC2188;g__HTCC	0.00	0.00	0.05	0.13	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Chromatiales;f__;g__	0.10	0.01	0.20	0.00	0.10	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__HTCC2188;f__HTCC2089;g__	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__;g__	0.00	0.09	0.10	0.03	0.01	0.00	0.01
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Coxiellaceae;g__	0.10	0.01	0.29	0.00	0.10	0.02	0.02
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Coxiellaceae;g__Aquicella	0.00	0.05	0.00	0.01	0.00	0.00	0.00

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Legionellales;f__Legionellaceae;g__Legionella	0.00	0.04	0.01	0.02	0.04	0.00	0.03
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Hahellaceae;g__Hahella	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Halomonadaceae;g__Halomonas	0.00	0.00	0.00	0.00	0.00	0.00	0.02
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas	0.00	0.01	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Thiotrichales;f__Piscirickettsiaceae;g__	0.20	0.23	0.68	0.00	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__	0.90	2.34	1.61	1.28	0.03	0.02	0.03
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Dokdonella	0.00	0.08	0.03	0.05	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Luteimonas	0.00	0.00	0.00	0.03	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Lysobacter	0.20	0.16	0.27	0.71	0.25	0.03	0.00
	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Thermomonas	0.20	0.47	0.52	0.45	0.00	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__TA18;o__CV90;f__;g__	0.00	0.14	0.01	0.00	0.03	0.00	0.00
	k__Bacteria;p__Proteobacteria;c__TA18;o__PHOS-HD29;f__;g__	0.10	0.19	0.20	0.00	0.15	0.00	0.04
	k__Bacteria;p__TM6;c__SJA-4;o__;f__;g__	0.00	0.00	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__TM7;c__SC3;o__;f__;g__	0.00	0.00	0.12	0.03	0.05	0.00	0.02
	k__Bacteria;p__TM7;c__TM7-1;o__;f__;g__	1.10	0.30	1.98	0.71	2.44	0.29	0.61
	k__Bacteria;p__TM7;c__TM7-3;o__;f__;g__	0.00	0.00	0.01	0.02	0.00	0.03	0.00
	k__Bacteria;p__TM7;c__TM7-3;o__I025;f__;g__	0.10	0.00	0.07	0.26	0.03	0.07	0.03
	k__Bacteria;p__Verrucomicrobia;c__Opitutae;o__Opitutales;f__Opitutaceae;g__Opitutus	0.00	0.05	0.07	0.02	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__	0.00	0.01	0.01	0.00	0.01	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Luteolibacter	0.10	0.25	0.22	0.34	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Verrucomicrobium	0.00	0.01	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Methylacidiphilae];o__S-BQ2-57;f__;g__	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__;g__	0.10	0.08	0.21	0.04	0.11	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__Ellin515;g__	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__Ellin517;g__	0.00	0.01	0.04	0.01	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Pedosphaerae];o__[Pedosphaerales];f__R4-41B;g__	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];Other	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];g__	0.10	0.30	0.09	0.02	0.05	0.00	0.04
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];g__Candidatus	0.00	0.00	0.00	0.00	0.01	0.01	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];g__Chthoniobacter	0.10	0.29	0.17	0.00	0.03	0.01	0.02

	Genera	Total	T1	T2	T3	T4	T5	T6
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];g__DA101	1.30	0.41	1.20	0.13	2.57	0.12	3.26
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];g__Ellin506	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	k__Bacteria;p__Verrucomicrobia;c__[Spartobacteria];o__[Chthoniobacterales];f__[Chthoniobacteraceae];g__heteroC45_4W	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	k__Bacteria;p__WPS-2;c__;o__;f__;g__	0.00	0.05	0.00	0.00	0.01	0.00	0.01
	k__Bacteria;p__WS3;c__PRR-12;o__Sediment-1;f__PRR-10;g__	0.00	0.04	0.00	0.00	0.00	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Deinococcaceae;g__CM44	0.00	0.00	0.00	0.00	0.02	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Deinococcaceae;g__Deinococcus	0.00	0.04	0.00	0.00	0.01	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Trueperaceae;g__	0.10	0.01	0.08	0.46	0.00	0.04	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Trueperaceae;g__B-42	0.10	0.00	0.00	0.78	0.00	0.00	0.00
	k__Bacteria;p__[Thermi];c__Deinococci;o__Deinococcales;f__Trueperaceae;g__Truepera	0.10	0.00	0.03	0.13	0.06	0.51	0.12

Appendix 6. Metabiodiverse Taxonomy Results

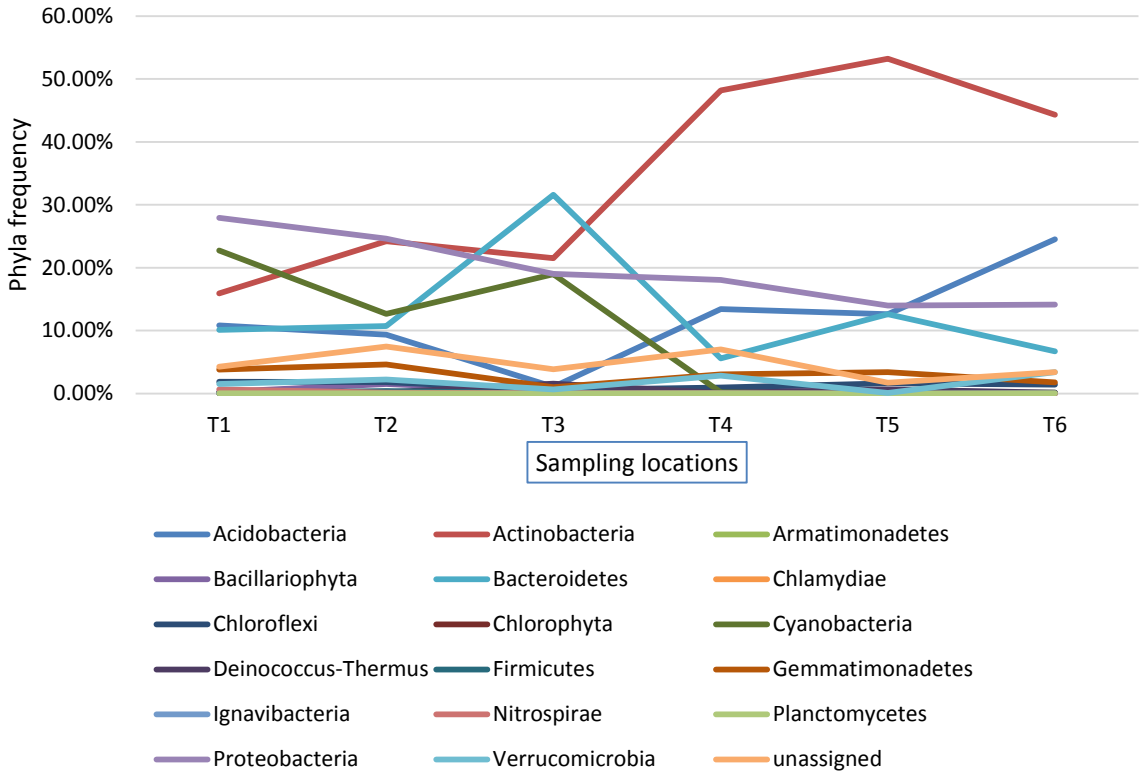


Figure 1. Phyla-level frequency per sampling site

Table 1. Detailed phyla-level assignments per sampling site

Phyla	T1	T2	T3	T4	T5	T6
Acidobacteria	10.80%	9.36%	1.06%	13.43%	12.60%	24.50%
Actinobacteria	15.89%	24.20%	21.48%	48.19%	53.23%	44.31%
Armatimonadetes	0.05%	0.09%	0.00%	0.06%	0.01%	0.05%
Bacillariophyta	0.21%	1.50%	0.34%	0.64%	0.07%	0.07%
Bacteroidetes	10.09%	10.74%	31.56%	5.55%	12.62%	6.67%
Chlamydiae	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%
Chloroflexi	1.84%	1.79%	0.55%	0.92%	1.60%	1.41%
Chlorophyta	0.05%	0.01%	0.00%	0.00%	0.00%	0.00%
Cyanobacteria	22.73%	12.65%	18.94%	0.17%	0.10%	0.13%
Deinococcus-Thermus	0.05%	0.17%	1.57%	0.09%	0.64%	0.14%
Firmicutes	0.10%	0.39%	0.04%	0.01%	0.00%	0.02%
Gemmatimonadetes	3.80%	4.63%	0.97%	3.04%	3.40%	1.76%
Ignavibacteria	0.01%	0.04%	0.01%	0.00%	0.00%	0.00%
Nitrospirae	0.63%	0.15%	0.02%	0.02%	0.00%	0.00%
Planctomycetes	0.04%	0.03%	0.00%	0.02%	0.00%	0.04%
Proteobacteria	27.90%	24.61%	18.99%	18.02%	13.96%	14.13%
Verrucomicrobia	1.52%	2.20%	0.64%	2.84%	0.10%	3.39%
unassigned	4.27%	7.44%	3.85%	6.98%	1.69%	3.39%

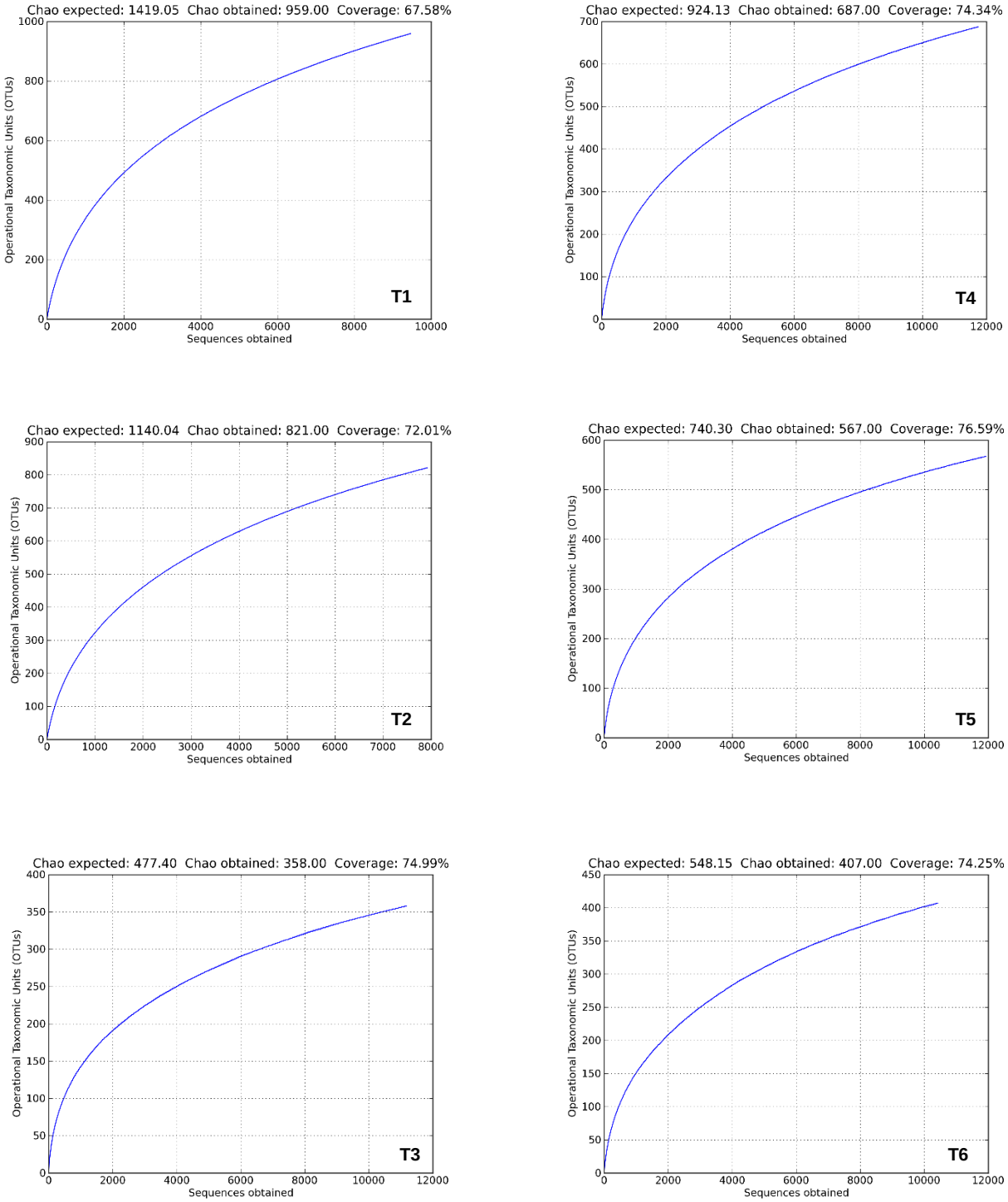


Figure 2. Rarefaction curves for the chao1 richness estimator. Observed richness is also depicted, in each graph.

Appendix 7. SILVAngs Taxonomy Results

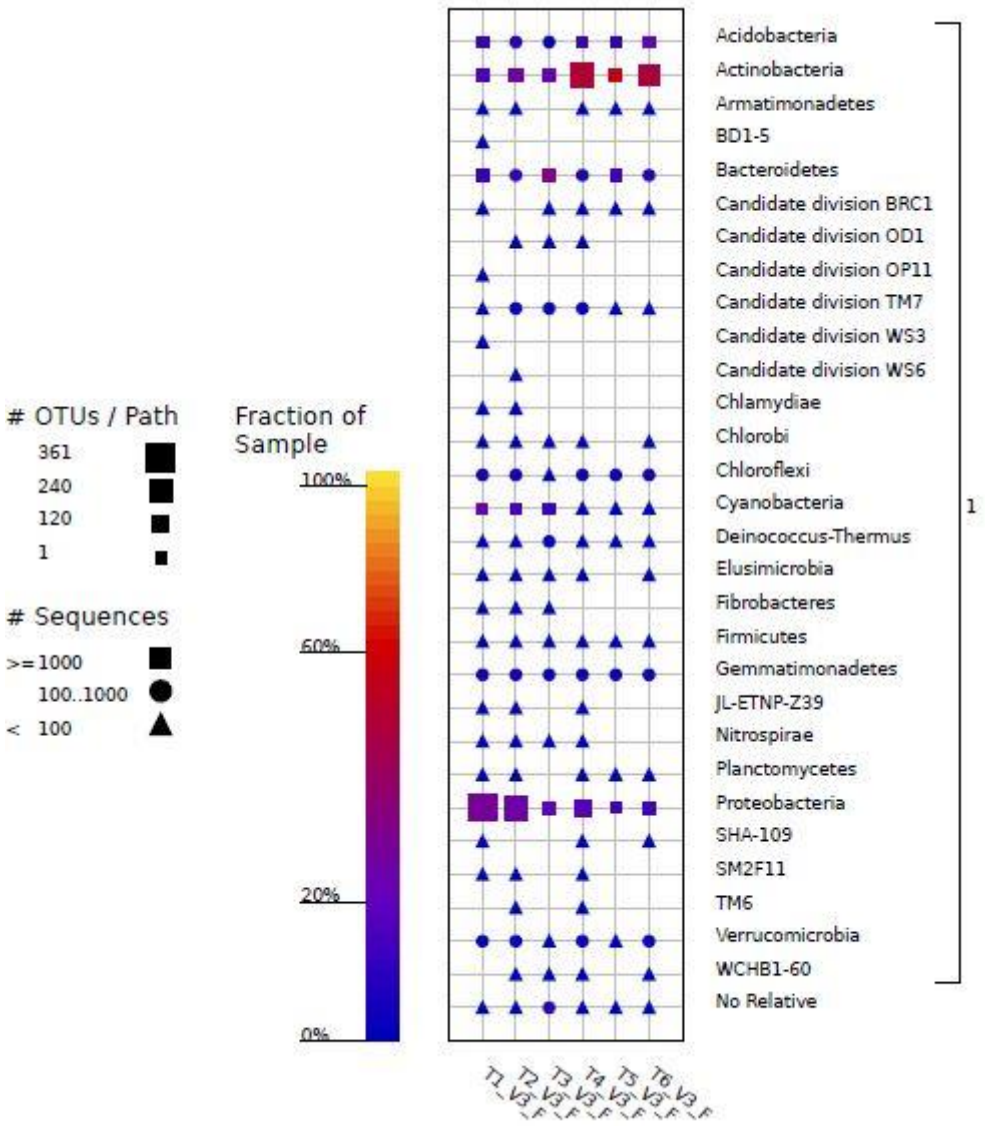


Figure 1. Phyla-level taxonomy report. The abundance (number of sequences) is given by a different shape, whereas the diversity (number of OTUs) is given by the size of the shape.

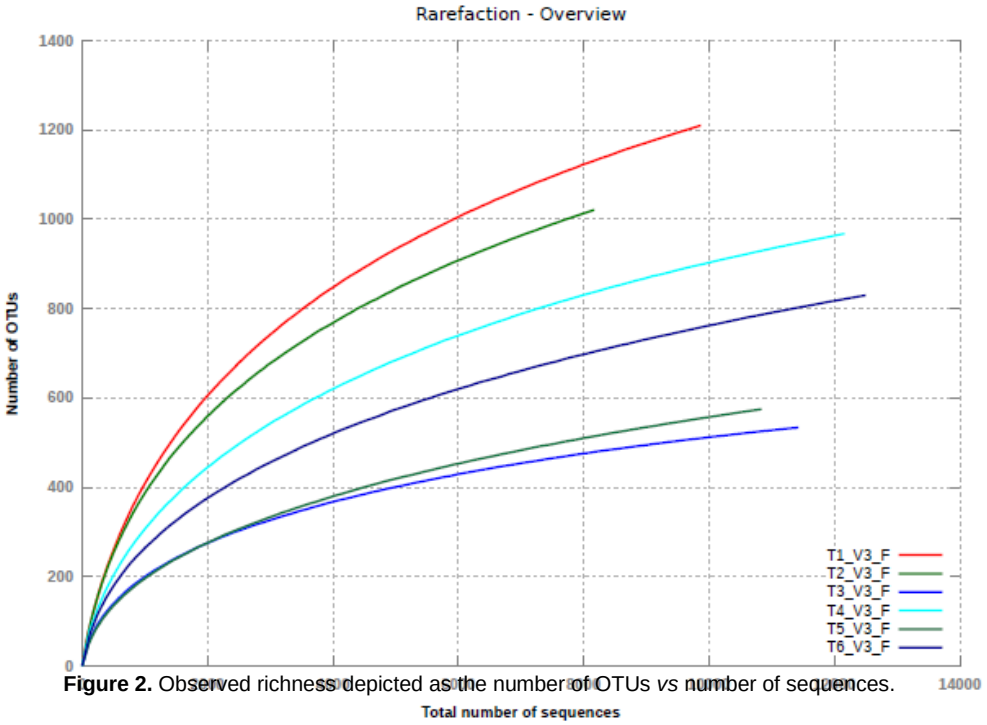


Figure 2. Observed richness depicted as the number of OTUs *vs* number of sequences.